



UNIVERSITA' DEGLI STUDI DI VERONA

DIPARTIMENTO DI ANGLISTICA, GERMANISTICA E SLAVISTICA

**DOTTORATO DI RICERCA IN LINGUISTICA**

CICLO XXI

***Valutare le competenze orali in italiano L2:  
variazione longitudinale e situazionale  
in apprendenti a livello avanzato***

S.S.D. L-LIN/02

Coordinatore: Prof.ssa Camilla Bettoni

Tutor: Prof.ssa Camilla Bettoni

Co-Tutor: Prof.ssa Ineke Vedder

Dottorando: Dott.ssa Stefania Ferrari

# Indice

<b>0</b>	<b>Prefazione</b>	<b>i</b>
<b>1</b>	<b>INTRODUZIONE</b>	<b>1</b>
1.1	<i>L'italiano L2 a scuola</i>	1
1.2	<i>L'ambito della ricerca</i>	4
1.3	<i>Gli obiettivi generali della ricerca</i>	7
<b>2</b>	<b>QUADRO TEORICO</b>	<b>9</b>
2.1	<i>Le competenze linguistiche</i>	10
2.1.1	La competenza	10
2.1.2	La proficiency, l'abilità comunicativa e l'abilità d'uso	14
2.2	<i>L'approccio acquisizionale</i>	25
2.2.1	La variazione	25
2.2.1.1	L'approccio sociolinguistico	27
2.2.1.2	L'approccio psicolinguistico	34
2.2.1.3	L'approccio emergentista	42
2.2.2	La misurazione della variazione	49
2.2.2.1	L'unità di produzione	51
2.2.2.2	Le misure analitiche	56
2.3	<i>L'approccio del testing</i>	71
2.3.1	La variazione	72
2.3.2	La misurazione della variazione	76
2.3.2.1	Le scale di proficiency	77
2.3.2.2	Il Quadro Comune Europeo di Riferimento	81

2.4	<i>L'interfaccia acquisizione-testing</i>	88
2.4.1	L'acquisizione e il testing	89
2.4.2	I costrutti, la misurazione e le scale di proficiency	92
2.5	<i>I task</i>	95
2.5.1	La definizione di task	96
2.5.2	I task per insegnamento, SLA e testing	99
<b>3</b>	<b>METODOLOGIA DI RICERCA</b>	<b>103</b>
3.1	<i>I soggetti</i>	104
3.2	<i>Le rilevazioni</i>	107
3.3	<i>I task</i>	109
3.4	<i>I dati</i>	110
3.5	<i>L'unità e le misure dell'analisi</i>	111
3.5.1	Le misure CAF	112
3.5.2	Le valutazioni QCER	132
<b>4</b>	<b>RISULTATI</b>	<b>139</b>
4.1	<i>La complessità sintattica CAF</i>	139
4.1.1	La variazione longitudinale	140
4.1.2	La variazione situazionale	142
4.1.3	La variazione individuale	147
4.1.4	Riassunto	186
4.2	<i>L'accuratezza CAF</i>	187
4.2.1	La variazione longitudinale	188
4.2.2	La variazione situazionale	189
4.2.3	La variazione individuale	191
4.2.4	Riassunto	203

4.3	<i>La fluenza CAF</i>	205
4.3.1	La variazione longitudinale	206
4.3.2	La variazione situazionale	207
4.3.3	La variazione individuale	209
4.3.4	Riassunto	213
4.4	<i>Complessità, accuratezza e fluenza CAF a confronto</i>	214
4.4.1	La variazione longitudinale	215
4.4.2	La variazione situazionale	216
4.4.3	La variazione individuale	219
4.4.4	Riassunto	224
4.5	<i>Le valutazioni QCER</i>	226
4.5.1	La variazione longitudinale	226
4.5.2	La variazione situazionale	228
4.5.3	Riassunto	231
4.6	<i>Misure CAF e valutazioni QCER a confronto</i>	232
4.6.1	La variazione longitudinale	232
4.6.2	La variazione situazionale	233
4.6.2	Riassunto	235
<b>5</b>	<b>CONCLUSIONE</b>	<b>237</b>
5.1	<i>Riassunto</i>	238
5.2	<i>Implicazioni per la ricerca, il testing e l'insegnamento</i>	245
	<b>Riferimenti bibliografici</b>	<b>249</b>
	<b>Appendici</b>	<b>279</b>
	Tabella 1, Task-based studies: pianificazione	279
	Tabella 2, Task-based studies: ripetizione	285
	Tabella 3, Task-based studies: caratteristiche dei task	287

## **Prefazione**

La valutazione è un tratto costitutivo dell'insegnamento, come lo sono la programmazione e la realizzazione del piano di lavoro attraverso le attività in classe. Sembra naturale: se si vuole definire la programmazione è necessario avere informazioni sullo stato delle competenze di partenza dell'allievo; per sapere come si sta procedendo con l'attività didattica, e eventualmente apportare dei miglioramenti, è importante avere strumenti di controllo e di monitoraggio in itinere; così come servono dispositivi per conoscere gli esiti alla fine del percorso, quando l'istituzione scolastica richiede di assegnare un valore all'apprendimento di ogni studente. La mia esperienza di insegnante e formatrice mi ha permesso di constatare che non esiste tema più delicato della valutazione nel mondo della scuola: ogni insegnante sembra interpretare diversamente il compito di valutare gli apprendimenti linguistici, e uno stesso studente è facilmente classificato come competente da un collega e di scarso rendimento da un altro. Testimoni di ciò le tante e variegata prove e griglie di valutazione incontrate in questi anni o le curiose affermazioni di qualche insegnante, tra cui la mia preferita è "il mio cinque è un otto". Relativamente all'insegnamento delle lingue, definire criteri e standard per descrivere i livelli

di partenza e registrare i progressi degli apprendenti richiede idee chiare su almeno due aspetti: quali caratteristiche hanno i diversi compiti comunicativi e quali tappe percorre lo sviluppo delle competenze linguistiche nel tempo. Da qui l'importanza di una stretta collaborazione tra chi si occupa di ricerca acquisizionale e di testing e chi opera direttamente nella scuola.

L'ingresso di studenti immigrati e i nuovi orientamenti curriculari, più orientati alla personalizzazione dei percorsi di apprendimento, nonché i risultati nazionali delle valutazioni del programma OCSE PISA, hanno riaperto negli ultimi tempi la discussione all'interno della scuola sul tema della valutazione. In ambito linguistico, sia per le lingue straniere sia per l'italiano L2, il Quadro Comune Europeo di Riferimento e il Portfolio Europeo delle Lingue godono di una certa diffusione. Da qui la motivazione per lo studio intrapreso con questa tesi, il cui obiettivo è appunto contribuire alla descrizione della natura e dello sviluppo delle competenze linguistiche in L2. Nonostante il vivace interesse per l'italiano L2, in letteratura infatti sono pochi gli studi che indagano la variazione longitudinale in apprendenti a livello avanzato inseriti nella scuola secondaria, e a mia conoscenza ancora nessuna ricerca verte sulla variazione situazionale legata a diverse attività comunicative. In questa tesi ho cercato quindi di affrontare questi temi coniugando le due prospettive dell'acquisizione e del testing, e tenendo sempre presenti le implicazioni pratiche sia per la didattica sia per la valutazione delle competenze linguistiche degli studenti immigrati inseriti nella scuola italiana.

Il lavoro è diviso in 5 capitoli. Il capitolo 1 introduce il tema oggetto di questo studio. Si presentano prima alcune questioni relative al tema dell'italiano L2 a scuola, con particolare riferimento al bisogno di strumenti di valutazione efficaci, poi si definisce l'ambito della ricerca. Questa è un'indagine della variazione longitudinale e situazionale della produzione orale di studentesse italiane e di apprendenti a livello avanzato impegnate in task monologici e interattivi, misurata secondo due approcci: le misure analitiche CAF e le valutazioni soggettive QCER. Il capitolo 2 fornisce l'ancoraggio teorico di questo studio discutendone la letteratura di base. Si inizia definendo

la competenza e la proficiency, e poi si presenta il tema della variazione e della sua misurazione secondo i due approcci dell'acquisizione e del testing, discutendo anche i punti di contatto tra i due ambiti. Si conclude presentando il task, strumento comune alle due prospettive. La prima parte del capitolo 3 è dedicata alla presentazione dettagliata degli obiettivi della ricerca, delle modalità di raccolta dati, delle informazioni relative ai partecipanti; la seconda parte alla presentazione delle misure impiegate. Il capitolo 4 riporta l'analisi, presentando prima i risultati della misurazione CAF, poi quelli delle valutazioni QCER, e infine discutendo le possibili interazioni tra i due strumenti. Nel capitolo conclusivo vengono sintetizzati i risultati e presentate alcune implicazioni per la ricerca acquisizionale, per il testing e per la didattica.

Questo studio non avrebbe potuto essere realizzato senza la preziosa disponibilità delle sei studentesse intervistate, l'efficiente collaborazione dell'I.P.S.I.A. Malpighi di San Giovanni in Persiceto (Bo) e del Multicentro Educativo Sergio Neri di Modena (M.E.MO). Inoltre, da tempo Gabriele Pallotti motiva e sostiene il mio interesse per la ricerca, Beatrice Iori stimola la mia passione per l'insegnamento. Più puntualmente i miei tutor accademici, Camilla Bettoni e Ineke Vedder, hanno seguito per tre anni l'evoluzione di questo studio con vigile intelligenza e costante attenzione; e Elena Nuzzo, Giovanna Masiero e Lucia di Lucca hanno condiviso le fatiche del fieldwork (e non solo). Infine alle scuole, agli insegnanti e soprattutto agli studenti con cui ho lavorato in questi anni devo lo stimolo, la passione e la curiosità che mi hanno permesso di realizzare di questo lavoro. A tutti va il mio più sincero ringraziamento.

# **CAPITOLO 1**

## **INTRODUZIONE**

### *1.1. L'italiano L2 a scuola*

La presenza in classe di alunni immigrati è un elemento strutturale del sistema scolastico italiano. I dati riportati dall'ultima indagine del ministero dell'istruzione (MIUR, 2008) stimano che i ragazzi immigrati seduti tra i banchi di scuola sono più di 570 mila, con un'incidenza del 6,4% sul totale della popolazione scolastica.

L'apprendimento/insegnamento dell'italiano come lingua seconda (L2 da ora in poi) rappresenta una componente essenziale del processo di integrazione degli alunni immigrati, in quanto la conoscenza della lingua è loro indispensabile per partecipare attivamente e con successo alla vita scolastica (e non). La loro presenza così numerosa ha favorito quindi per l'apprendimento/insegnamento dell'italiano L2 e stimolato il bisogno di intraprendere ricerche sistematiche per capirne meglio i processi.

La scolarizzazione degli alunni immigrati è caratterizzata da almeno due nodi problematici: la scelta tra i diversi indirizzi al termine della scuola



dell'obbligo e il successo scolastico. Riguardo al primo problema, gli alunni immigrati sembrano preferire gli istituti professionali in percentuale doppia rispetto agli italiani (40,6% vs 19,9% secondo MPI, 2006). Questo dato suggerisce il rischio di segregazione o di discriminazione. Riguardo al secondo problema, il confronto tra gli studenti italiani e immigrati rileva come l'esito scolastico di questi ultimi sia costantemente inferiore in tutti gli ordini di scuola. Nelle scuole secondarie di II grado oltre un quarto degli allievi stranieri non viene promosso, circa il doppio degli studenti italiani. Negli altri ordini di scuola, il divario è rispettivamente del 3,2% nella scuola primaria e del 7,9% nella secondaria di I grado (MPI, 2006). Nonostante questi dati nazionali nascondano situazioni diverse e frastagliate tra tipologie di scuole, province e regioni, è evidente come non ci si possa oggi limitare a organizzare misure compensatorie di carattere speciale, ma si debba invece richiedere a insegnanti e ricercatori uno sforzo congiunto per l'elaborazione di metodologie d'intervento efficaci.

La misurazione e valutazione delle competenze linguistiche degli alunni è prerequisite essenziale per organizzare politiche educative adeguate e attivare 'buone pratiche' di educazione linguistica. Per intervenire nel modo più efficace possibile gli insegnanti hanno bisogno di strumenti che permettano loro di diagnosticare i livelli linguistici di partenza e quelli di arrivo degli allievi, determinarne il progresso, controllare il raggiungimento degli obiettivi, selezionare, definire e identificare gli standard di apprendimento. Le questioni poste dalla presenza degli alunni immigrati non possono che essere uno stimolo per mantenere attiva nella scuola una riflessione più generale sulla qualità della valutazione e degli apprendimenti linguistici. Come sottolinea il documento ministeriale *La via italiana per l'intercultura e l'integrazione degli alunni stranieri*, "[L]a presenza dei minori stranieri funziona in realtà da evidenziatore di sfide che comunque la scuola italiana dovrebbe affrontare anche in assenza di stranieri" (Osservatorio Nazionale, 2007: 6).

La scuola di oggi sta cambiando. La legge sull'autonomia, il Piano dell'Offerta Formativa, l'introduzione del Portfolio delle Competenze e dei

Piani di Studio Personalizzati sono tra le tappe più significative del cambiamento in atto. Elemento comune delle diverse riforme scolastiche che si sono succedute negli ultimi anni sembra essere proprio l'aumentata attenzione per l'alunno, favorevole a forme di insegnamento/apprendimento più adatte alle diverse esigenze degli apprendenti. Nella stessa direzione si è mossa la normativa relativa all'accoglienza e integrazione degli alunni stranieri, anch'essa favorevole alla flessibilità dei programmi e alla personalizzazione della valutazione. La normativa affida quindi un ruolo chiave alle singole istituzioni scolastiche autonome, rafforzando la responsabilità dei docenti nella definizione del necessario adattamento dei programmi e della valutazione, così come si evince dalle *Linee guida per l'integrazione*:

Il riferimento più congruo a questo tema [la valutazione] lo si ritrova nell'art. 45, comma 4, del DPR n 394 del 31 agosto 1999 che così recita "il collegio dei docenti definisce, in relazione al livello di competenza dei singoli alunni stranieri, il necessario adattamento dei programmi di insegnamento ...". Benché la norma non accenni alla valutazione, ne consegue che il possibile adattamento dei programmi per i singoli alunni comporti un adattamento della valutazione, anche in considerazione degli orientamenti generali su questo tema, espressi in circolari e direttive, che sottolineano fortemente l'attenzione ai percorsi personali degli alunni. (MIP, 2003: 16).

La valutazione è un tratto costitutivo dell'insegnare, come lo sono la programmazione e la conseguente realizzazione del piano di lavoro attraverso le attività in classe. Eppure lavorando con gli insegnanti emerge spesso un sentimento di disagio nei confronti di questa attività, in particolare quando coinvolge gli alunni immigrati. Questo disagio sembra nascere da almeno due elementi di incongruenza nell'operato che viene loro richiesto. L'imprescindibile esigenza di prestare maggiore attenzione al processo di apprendimento dei singoli alunni appare incongruente, in primo luogo, con la richiesta delle istituzioni di prevedere un numero preciso di prove, e di tradurne l'esito in voto o giudizio sintetico, e in secondo luogo, con la richiesta di portare tutti gli allievi a un pari livello di competenza e al raggiungimento di standard comuni.

Da questo quadro risulta chiaro come la valutazione costituisca un problema delicato per gli operatori della scuola, cui viene richiesto di

sviluppare strumenti utili per registrare i progressi delle competenze in italiano L2 degli alunni immigrati. Il presente studio si propone di contribuire a risolvere questo problema, discutendo la natura, lo sviluppo e la misurazione della proficiency linguistica.

## *1.2. L'ambito della ricerca*

La linguistica acquisizionale e il testing si occupano entrambi delle lingue seconde, e dunque appartengono al medesimo ambito di studio, ma costituiscono due campi diversi. Le ricerche acquisizionali si concentrano principalmente sui processi dell'apprendimento e sui fattori che lo favoriscono o ostacolano con l'intento di spiegarli. Le ricerche sul testing invece si concentrano sugli elementi e sulle strategie che compongono l'abilità linguistica, elaborando teorie relative alla performance in grado di spiegarne la variazione durante i test e di controllarne la corrispondenza con l'uso della lingua in contesti reali. La differenza di sguardo che i due campi di ricerca assumono verso la L2 si riflette anche negli approcci metodologici che impiegano. Mentre le ricerche acquisizionali usano metodologie etnografiche o protocolli sperimentali per l'analisi linguistica delle produzioni interlinguistiche, le ricerche sul testing si affidano principalmente a correlazioni statistiche elaborate sui risultati ottenuti nei test.

Queste differenze di sguardo e di metodologia hanno finora tenuto divisi i due campi. Eppure già da tempo alcuni lavori (cfr. Bachman, 1988; Bachman & Cohen, 1998a; Hulstijn & Schoonen, 2006; e § 2.4) riconoscono le potenzialità di uno scambio proficuo tra la linguistica acquisizionale e le ricerche sul testing linguistico. Chi si occupa di testing infatti può ritrovare nella letteratura acquisizionale elementi importanti per una migliore definizione degli elementi linguistici testabili ai diversi livelli di proficiency. D'altro canto, chi si occupa di ricerca acquisizionale può imparare dalla letteratura relativa al testing l'importanza di una più rigorosa definizione

operativa dei costrutti teorici e della verifica di validità e affidabilità degli strumenti di elicitazione e misurazione. Un punto di contatto tra ricerche acquisizionali e testing è sicuramente rappresentato dal tema della variabilità delle produzioni. E in questa prospettiva si inserisce il presente studio, che intende analizzare la variabilità integrando appunto l'approccio del testing con quello dell'apprendimento della L2.

In letteratura è ormai assodato che le produzioni degli apprendenti, così come quelle dei parlanti nativi, siano variabili (cfr. § 2.2.1). Gli studi acquisizionali distinguono almeno due tipi di variabilità dell'interlingua, dando per scontata la ben nota variabilità individuale. Un primo tipo di variabilità, definita da Corder (1981) come *vertical continuum*, è intrinseca all'interlingua, che come sistema autonomo e transitorio è di per sé variabile. In questo senso l'attenzione di chi la studia viene posta verso l'alternanza tra forme native e forme non native, o tra due o più forme non native per esprimere una determinata nozione che nella lingua target viene espressa con una sola forma. Siccome la variazione riscontrata nell'interlingua è segno che l'apprendente sta acquisendo la forma invariabile della lingua target, l'apprendimento si esprimerà con una riduzione di tale variazione. Un secondo tipo di variabilità, chiamata *horizontal continuum* da Corder (1981), riguarda gli aspetti della lingua target in cui i parlanti nativi presentano variazione sociolinguistica e situazionale. Il progresso in questo caso è visto come acquisizione dell'intera gamma delle varianti native, usate con la medesima frequenza dei parlanti nativi, nel medesimo rispetto dei fattori linguistici ed extra-linguistici che le vincolano. Pertanto l'apprendimento si esprimerà con l'aumento di tale variazione, non con la riduzione: se la variabilità orizzontale è parte della competenza del parlante nativo, essa deve essere acquisita dall'apprendente.

Il tema della variabilità è centrale anche nel campo della valutazione. Se le produzioni di un apprendente variano è essenziale determinare quale tipo di performance è rappresentativa delle abilità che si vogliono testare. Come spiega Davies (1990: 137) "the fundamental argument/debit in language testing over the last 25 years has been basically about the meaning/realization of

language behaviour”. Poiché possiamo determinare quello che un apprendente sa esclusivamente da quello che fa (cfr. Ellis, 2001), qualsiasi forma di testing è *performance testing*. Il compito dei ricercatori è dunque quello di stabilire quale tipo di performance testare.

Se allora la variabilità dell’interlingua è centrale in entrambi i campi della linguistica acquisizionale e del testing, diventa importante rilevarla con una metodologia adeguata. Oggi il task sembra essere il mezzo più opportuno (cfr. § 2.5). Il task, riprendendo la definizione di Skehan (1996b), si intende come un’attività linguistica in cui l’attenzione del parlante è primariamente volta al significato piuttosto che alla forma di quello che dice, poiché l’esito del compito viene valutato sul risultato raggiunto. In questo modo il tutto – anche se compiuto per scopi di ricerca, di testing o di didattica – ha una qualche relazione con il mondo reale. A partire dagli anni ’70 il task è stato proposto in didattica come valida alternativa per l’organizzazione del syllabo. I task pedagogici hanno infatti il vantaggio di fornire agli apprendenti occasioni sia per praticare forme già conosciute sia per apprenderne di nuove, favorendo un uso naturale della lingua e offrendo contemporaneamente specifici momenti di *focus on form*. Per costruire un syllabo efficace è necessario però avere a disposizione un quadro di riferimento che permetta di concettualizzare la performance richiesta dai vari task, di selezionarli e di graduarli sulla base delle loro caratteristiche, e infine di implementarli per massimizzare le occasioni di *focus on form*. Di conseguenza in ambito acquisizionale la ricerca si è concentrata su una serie di aspetti dei vari task tra cui l’analisi della loro difficoltà in relazione alle competenze e alle caratteristiche degli apprendenti (cfr. 2.2.1.2). Queste ricerche osservano in genere la variazione della lingua prodotta secondo tre dimensioni (cfr. Skehan, 1996a: 22): (a) la complessità sintattica, ossia la lunghezza e il grado di subordinazione degli enunciati; (b) l’accuratezza, ossia il grado di conformità alla norma della lingua target; e (c) la fluenza, ossia la capacità di evitare eccessive pause o esitazioni. Secondo Skehan (1998) e Robinson (2001a), queste tre dimensioni della performance sono indipendenti l’una dall’altra e vengono influenzate dalle caratteristiche

del task e dal conseguente tipo di processing che attivano nell'apprendente. Queste tre dimensioni sono centrali per il mio studio, e mi riferirò ad esse usando l'acronimo CAF proposto da Alex Housen, Folkert Kuiken e Ineke Vedder nei loro vari lavori che verranno citati specificamente nel corso della tesi.

Nell'ambito del testing è in corso un dibattito sull'applicabilità del task quale strumento per aumentare la validità ecologica delle prove e di conseguenza la loro generalizzabilità a contesti di vita reale (cfr. *Language Testing*, 2002). Questo è coerente sia con l'approccio comunicativo alla didattica della L2 sia con l'approccio alla valutazione basato sulla competenza d'azione, così come descritta nel Quadro Comune Europeo di Riferimento per le lingue (QCER da ora in poi) (Consiglio d'Europa, 2001). Anche in questo caso una maggiore comprensione della variabilità nella performance elicitata tramite i diversi task e del rapporto che intercorre tra abilità e performance in un task pare essenziale, in quanto fornisce una base per selezionare i task secondo le difficoltà che pongono ai candidati (McNamara, 1996; Brown *et al.*, 2002). In questo ambito la valutazione della performance viene solitamente realizzata a partire da scale di descrittori della proficiency.

### *1.3. Gli obiettivi generali della ricerca*

Nel presente lavoro ci si propone da un lato di investigare in prospettiva acquisizionale la variabilità situazionale e longitudinale dell'interlingua di apprendenti avanzati impegnati nello svolgimento di task comunicativi, e dall'altro di verificare sperimentalmente con prospettiva legata al testing le relazioni che intercorrono tra due tipi di misurazione: quella analitica CAF che rileva la complessità, l'accuratezza e la fluenza della produzione orale, impiegata nei cosiddetti *developmental index studies* di stampo acquisizionale, e quella soggettiva basata sui descrittori del QCER, maggiormente legata all'ambito del testing.

Questa ricerca ha un carattere essenzialmente esplorativo, nel senso che si pone l'obiettivo di rispondere ad alcune domande-guida facendo scaturire eventuali ipotesi piuttosto che di verificare ipotesi già formulate e/o di convalidare teorie specifiche. In particolare attraverso l'analisi quantitativa e qualitativa del parlato di apprendenti a livello avanzato confrontato con quello di parlanti nativi si cercherà di rispondere alle seguenti domande:

- Come variano complessità, accuratezza e fluenza nel corso del tempo man mano che progredisce l'interlingua? Come variano con il variare dell'attività comunicativa per gli apprendenti rispetto ai parlanti nativi? Come si sviluppa il percorso individuale di ciascun soggetto? Quali sono le interazioni tra queste tre dimensioni dell'interlingua?
- I descrittori del QCER sono efficaci nella valutazione delle competenze orali di apprendenti di italiano L2 di livello avanzato? In particolare sono sensibili allo sviluppo longitudinale e alla variazione tra task?
- Quali relazioni intercorrono tra le misure analitiche CAF e le valutazioni soggettive della proficiency basate sul QCER?

## **CAPITOLO 2**

### **QUADRO TEORICO**

Il capitolo è dedicato alla presentazione dei concetti teorici, dei costrutti analitici e delle misure applicative che forniscono la base per questo studio. Per valutare le competenze linguistiche è necessario non solo capire che cosa siano, ma anche definirne operativamente i costrutti e individuare strumenti di misurazione efficaci. Pertanto verranno dapprima introdotte le nozioni di competenza e proficiency, e discusse le diverse interpretazioni date a questi termini in letteratura (§ 2.1). Quindi si affronterà il tema della variabilità delle produzioni interlinguistiche: se la performance è variabile, sarà necessario stabilire cosa vale la pena osservare e con quali strumenti. A questo proposito si considereranno due prospettive: quella acquisizionale (§ 2.2) e quella relativa al testing linguistico (§ 2.3), riportando la discussione teorica sul ruolo che la variazione assume in ciascuna e le modalità di valutazione tipiche dei due ambiti: le misure CAF e le valutazioni QCER. Si discuteranno poi i punti di contatto tra gli studi acquisizionali e la ricerca sul testing evidenziando i vantaggi di una maggiore interazione tra i due ambiti (§ 2.4). Infine si presenterà il task (§ 2.5), quale strumento comune alle due prospettive.



## 2.1. *Le competenze linguistiche*

Linguistica acquisizionale e ricerca sul testing condividono l'attenzione per le competenze linguistiche: mentre la prima però si occupa della descrizione dell'interlingua dell'apprendente e del suo sviluppo, la seconda si concentra sulla definizione dell'abilità linguistica. Dal punto di vista acquisizionale è importante capire i meccanismi e i fattori che entrano in gioco nello sviluppo delle competenze linguistiche, mentre dal punto di vista del testing è fondamentale avere una chiara definizione del costrutto dell'abilità linguistica che aiuti a spiegare cos'è la proficiency e come questa si relazioni ai vari test.

Definire in modo teoricamente valido e operativamente utile l'abilità linguistica si è rivelato però un compito estremamente complesso. Gli stessi ricercatori faticano a trovare un accordo e in letteratura l'interpretazione di termini quali competenza, proficiency o performance rimane spesso confusa. I prossimi paragrafi inquadrano i problemi più salienti e riassumono i termini del dibattito. Si conclude introducendo il quadro di riferimento impiegato in questo studio.

### 2.1.1. *La competenza*

Il termine competenza è sicuramente tra i più controversi nel campo della linguistica generale e della linguistica applicata. La confusione è dovuta al fatto che i vari autori usano questo termine riferendosi a nozioni sostanzialmente diverse.

Chomsky interpreta la competenza come una proprietà statica e assoluta dell'individuo. Nel suo testo *Aspects of the theory of syntax* definisce così la nozione di competenza, opponendola a quella di performance:

Linguistic theory is primarily concerned with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual *performance*.

[...] We thus make a fundamental distinction between *competence* (the speaker-hearer's knowledge of the language) and *performance*, the actual use of language in concrete situations. (Chomsky, 1965: 3-4)

La competenza in questo caso è dunque vista come uno stato mentale astratto che esclude l'abilità linguistica "it does not seem to me to be quite accurate to take 'knowledge of English' to be a capacity or ability, though it enters into the capacity or ability exercised in language use" (Chomsky, 1975: 23). Quello che interessa al linguista è il prodotto e non il processo, come chiarisce Taylor (1988: 153): l'idea di competenza di Chomsky non ha nulla da dire né rispetto all'abilità a impiegare la rappresentazione mentale nella comunicazione, né rispetto a come i parlanti usano o apprendono tale competenza.

L'idea di competenza intesa come conoscenza di un parlante ideale con un controllo perfetto della lingua in questione è stata presto criticata da chi aveva un approccio sociolinguistico allo studio del linguaggio e vedeva la distinzione effettuata da Chomsky come riduttiva, in quanto confinava tutti i casi di variabilità alla performance, escludendoli così dall'oggetto di indagine. Particolarmente influente in questo senso è il lavoro condotto da Hymes (1971; 1972), che, criticando l'attribuzione di tutti i fattori situazionali e socioculturali alla sola performance, sottolinea "the distorting effect of using the one term 'performance' for two distinct things: a theory of performance and a theory of language use" (1971: 11). In altre parole Hymes ritiene che alcuni aspetti di ciò che Chomsky relega alla performance siano sistematici, descrivibili sotto forma di regole, e pertanto da includere in un modello di competenza, che egli definisce appunto 'competenza comunicativa'. Lo stesso Chomsky concorderà con queste osservazioni, aggiungendo alla competenza grammaticale la competenza pragmatica, ossia la capacità di far uso della conoscenza definita come competenza grammaticale (Chomsky, 1980: 59).

Rispetto a Chomsky però, con l'introduzione della nozione di competenza comunicativa Hymes cambia profondamente il concetto di competenza, espandendolo in due direzioni. Innanzitutto va oltre la grammatica includendo "what is feasible, appropriate and actually performed". Poi

introduce nella nozione di competenza anche la dimensione sociale del linguaggio, includendo nel quadro l'abilità d'uso: "competence is understood to be dependent on two things: (tacit) knowledge and (ability for) use" (Hymes, 1971: 16; 1972: 282).

Nemmeno l'approccio adottato da Hymes è esente da critiche. Taylor (1988: 156) per esempio sostiene che Hymes, includendo nella definizione di competenza la nozione di abilità, avrebbe trasformato il concetto di competenza a tal punto da avergli fatto perdere un significato preciso:

For Hymes the general term *competence* covers a number of different elements, ranging from grammatical competence on the one hand to sociolinguistic competence on the other. (It is to be noted that Hymes's 'sociolinguistic competence' is much wider than Chomsky's 'pragmatic competence'.) The term applies then to something that Chomsky would see as biologically based (grammatical competence) and at the same time to something much more socially based (sociolinguistic competence). The former is purely individual, the latter is mainly social. The former concerns form; the latter concerns function. The former characterizes a state; the latter involves processes.

Canale & Swain (1980) e Canale (1983) riprendendo il lavoro di Hymes sviluppano il concetto di competenza comunicativa con particolare riferimento alla L2: la competenza non è costituita esclusivamente dalla conoscenza della grammatica della lingua, ma in modo anche più determinante dal modo in cui l'apprendente utilizza tale abilità per comunicare nei diversi contesti. La competenza comunicativa è pertanto costituita da quattro componenti: la competenza grammaticale, la competenza sociolinguistica, la competenza discorsiva e la competenza strategica. La competenza grammaticale è data dalla conoscenza del codice linguistico, e include pertanto le regole morfosintattiche, il lessico e la pronuncia. La competenza sociolinguistica riguarda la selezione delle forme linguistiche adatte a un determinato contesto e un determinato scopo. La competenza discorsiva è l'abilità che permette di scegliere forme e significati utili a organizzare coerentemente un testo o un messaggio. La competenza strategica permette infine di compensare la mancata conoscenza di regole della lingua in questione. Canale & Swain (1980) usano il termine competenza strategica per indicare le modalità con cui il parlante risolve un problema comunicativo: la competenza strategica è dunque per

definizione compensatoria, si attiva quando le altre competenze non ci sono o non sono in quel momento disponibili.

Competenza linguistica e competenza comunicativa indicano dunque due nozioni sostanzialmente diverse, una più restrittiva che esclude l'abilità d'uso e una più ampia che la include. Ellis (1994) ad esempio esplicita la distinzione sottolineando come la competenza sia "a language user's underlying knowledge of language" (p. 697) e la competenza comunicativa "the knowledge that users of a language have internalized to enable them to understand and produce messages in the language" (p. 696). Wiemann & Backlund (1980: 190) affrontano la definizione di competenza comunicativa da un punto di vista didattico, precisando come non possa che esservi un legame molto stretto tra abilità e competenza in quanto "*proficiency skills [...] are what is required for the manifestation of communicative competence*". D'altronde già Bruner (1973) includeva gli *skill* nella sua definizione di competenza sostenendo che quest'ultima è l'obiettivo dell'intervento didattico e pertanto può essere appresa: "what is learned is competence, not particular *performance*" (Bruner, 1973: 111).

Ricapitolando, possiamo osservare come la confusione che caratterizza la definizione di competenza sia legata all'ambito entro il quale le diverse definizioni sono nate: da un punto di vista strettamente linguistico infatti la competenza è vista come proprietà statica e assoluta, dal punto di vista educativo o psicologico, viene descritta come proprietà relativa e per definizione dinamica. Come riconosce Taylor (1988: 165-166) la confusione è dovuta al fatto che la definizione di competenza in senso chomskiano non sia adatta a far fronte alla variazione e al mutamento tipici dell'apprendimento. Quando come in questo studio l'obiettivo è indagare lo sviluppo linguistico, sia esso in L1 o in L2, diventa essenziale prendere in considerazione i processi ed essere pertanto in grado di distinguere "between what a speaker knows and what he does, between his knowledge and his proficiency in using that knowledge, between what is individual and what is social" (Taylor, 1988: 166). A questo proposito, rifacendoci a Taylor (1988: 166), in questo studio per

competenza e performance si intenderanno le nozioni chomskiane e si utilizzerà proficiency per indicare la capacità di utilizzare la competenza: “*Competence* can be regarded as a static concept, having to do with structure, state, or form, whereas proficiency is essentially a dynamic concept, having to do with process and function”. I principali modelli utilizzati per descrivere la proficiency saranno illustrati nella sezione che segue.

#### 2.1.2. La proficiency, l’abilità comunicativa e l’abilità d’uso

Il termine proficiency è strettamente legato alla letteratura relativa al testing linguistico, e come per la nozione di competenza la sua definizione è stata oggetto di dibattito. Morrow (1979/81: 18) ad esempio ne propone una definizione intuitiva quanto vaga: “how successful the candidate is likely to be as a user of the language in some general sense”. Altrettanto vaga e ingannevolmente semplice è la definizione di Ingram (1985: 220): “what is meant when we say someone is proficient in a language is that that person can do certain things in that language”. Altri autori ritengono la proficiency definibile in termini più restrittivi e tecnici come ciò che viene misurato dai test (Vollmer, 1981; Pienemann, 1985).

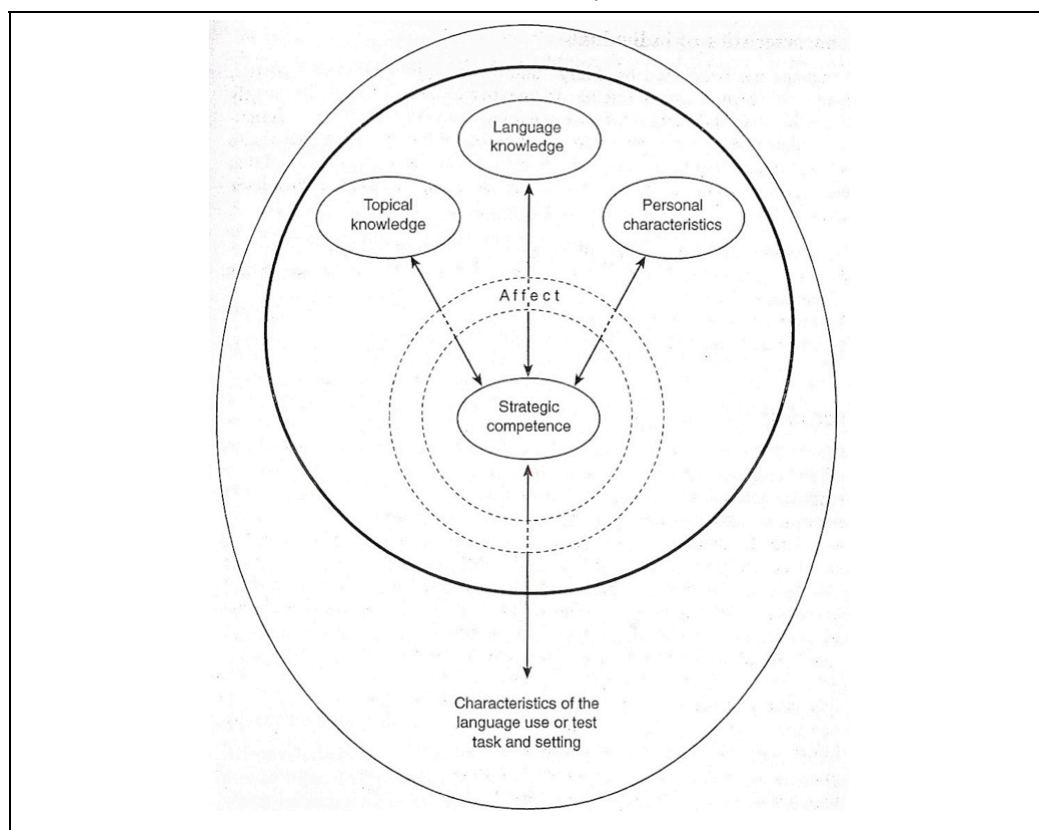
Stern (1983: 341) offre una definizione più precisa di proficiency descrivendola come “the actual performance of given individual learners or groups of learners”. Essa pertanto richiede una serie di diverse conoscenze e abilità (p. 346): (a) una conoscenza intuitiva delle forme della lingua; (b) una conoscenza intuitiva dei significati linguistici, cognitivi, affettivi e socioculturali espressi dalle forme linguistiche; (c) la capacità di utilizzare la lingua con attenzione massima al significato e minima alla forma; (d) la creatività dell’uso linguistico. Come evidenzia anche Taylor (1988: 163) Stern include nell’idea di proficiency sia la nozione statica di competenza elaborata da Chomsky, sia la definizione di abilità d’uso introdotta da Hymes. Stern (1983: 357) sottolinea inoltre come la proficiency possa essere indagata

secondo due prospettive: una che mira a stabilirne i livelli, in una progressione dal livello base a quello quasi-nativo, l'altra che si concentra invece sulle componenti della proficiency stessa.

Diversi studi, in particolare nell'area della didattica e del testing, si sono occupati della definizione delle componenti della proficiency linguistica. In una fase iniziale, sulla base delle ricerche condotte da Oller (1979; 1983) e la conseguente elaborazione della sua *Unitary Trait Hypothesis*, si riteneva che la proficiency fosse un costrutto unitario. Questa proposta subì diverse critiche, in particolare rispetto alle procedure statistiche impiegate. Vollmer & Sang (1983) ad esempio, rianalizzando i dati di Oller, ne contraddicono le conclusioni e sostengono piuttosto l'idea della proficiency come costituita da diverse componenti. Nello stesso periodo Bachman & Palmer (1981), utilizzando una *Multitrait-Multimethod Matrix*, dimostrano come la proficiency nel parlato e nella lettura siano indipendenti l'una dall'altra, e dunque anche loro sostengono l'idea della proficiency come multi-componenziale.

Il modello più influente nell'ambito del testing è certamente quello proposto da Bachman & Palmer (1996), raffinamento ulteriore del modello proposto in Bachman (1990). Questi autori non utilizzano il termine proficiency, ma preferiscono piuttosto l'espressione *Communicative Language Ability* (CLA da ora in poi). Il loro modello differisce da quelli che lo hanno preceduto per le seguenti caratteristiche: in primo luogo è un modello operativo, utile a mostrare il funzionamento di un test; in secondo luogo la CLA è solo un aspetto del modello tra altri che possono influire sulla performance, quali il metodo di valutazione o le caratteristiche dell'apprendente. Infine, questo modello assegna un ruolo centrale alla competenza strategica, intesa come autonoma dalla conoscenza linguistica.

**Schema 1. Some components of language use and language test performance, da Bachman & Palmer, 1996: 63**



Secondo Bachman (1990: 84) la CLA può essere descritta come “consisting of both knowledge, or competence, and the capacity for implementing, or executing that competence in appropriate, contextualized communicative language use”. L’uso linguistico coinvolge dunque una serie di interazioni complesse tra le caratteristiche individuali dell’apprendente da un lato e le caratteristiche della situazione comunicativa o di valutazione dall’altro. La CLA viene pertanto inquadrata all’interno di un quadro interazionale dell’uso linguistico (cfr. schema 1):

[the interactional framework] focuses on the interactions amongs areas of language ability (language knowledge and strategic competence, or metacognitive strategies), topical knowledge, and affective schemata, on the one hand, and how these interact with the characteristics of the language use situation, or test task, on the other. (Bachman & Palmer, 1996: 62)

Il modello proposto in Bachman & Palmer (1996) è costituito da 4 componenti: la conoscenza linguistica e la competenza strategica, la topical knowledge e le caratteristiche personali.

**Schema 2. Areas of language knowledge,  
da Bachman & Palmer, 1996: 68**

organizational knowledge (how utterances or sentences and texts are organised)
grammatical knowledge (how individual utterances or sentences are organised)
knowledge of vocabulary knowledge of syntax knowledge of phonology/graphology
textual knowledge (how utterances or sentences are organized to form texts)
knowledge of cohesion knowledge of rethorical or conversational organization
pragmatic knowledge (how utterances or sentences and texts are related to the communicative goals of the language users and to the features of the language use setting)
functional knowledge (how utterances or sentences and texts are related to the communicative goals of the language users)
knowledge of ideational functions knowledge of manipulative functions knowledge of heuristic functions knowledge of imaginative functions
sociolinguistic knowledge (how utterances or sentences and texts are related to features of the language use setting)
knowledge of dialects/varieties knowledge of registers knowledge of natural or idiomatic expressions knowledge of cultural references and figures of speech

Come mostra lo schema 2, la conoscenza linguistica a sua volta è suddivisa in due categorie, conoscenza organizzativa e conoscenza pragmatica. La conoscenza organizzativa coinvolge il controllo formale delle strutture linguistiche necessarie per produrre e comprendere frasi grammaticalmente



accettabili e per organizzare frasi e enunciati in testi scritti e orali; essa comprende pertanto due componenti: la conoscenza grammaticale e la conoscenza testuale. La conoscenza pragmatica guida l'interpretazione del discorso mettendo in relazione frasi ed enunciati con il loro significato, le intenzioni del parlante e le caratteristiche del contesto comunicativo; essa comprende pertanto due componenti: la conoscenza funzionale e la conoscenza sociolinguistica.

La componente competenza strategica è descritta invece come composta da “higher order executive processes that provide a cognitive management function in language use, as well as in other cognitive activities” (Bachman & Palmer, 1996: 70). A differenza dei modelli precedenti, la competenza strategica non è vista come abilità compensatoria, attivata per colmare lacune, ma ha un ruolo centrale: essa ha una funzione di mediazione tra intenzione comunicativa, competenze linguistiche, conoscenza del mondo e contesto situazionale. Tre sono le aree in cui operano le componenti metacognitive della competenza strategica: definizione degli obiettivi, valutazione e pianificazione. Lo schema 3 riporta le aree dell'uso della competenza strategica.

La CLA comprende dunque sia la proficiency linguistica, intesa come capacità d'uso della lingua, sia la proficiency comunicativa, che include la conoscenza del mondo e delle strategie necessarie per applicare la proficiency linguistica alle diverse situazioni. Come nota McNamara (1995: 172) la novità del modello di Bachman (1990) consiste nel riconoscere il ruolo degli aspetti della capacità d'uso delle competenze linguistiche ridefinendo il rapporto tra competenza e performance in modo dinamico.

**Schema 3. Areas of metacognitive strategy use,  
da Bachman & Palmer, 1996: 71**

goal setting
(deciding what one is going to do)
identifying the test tasks
choosing one or more tasks from a set of possible tasks (sometimes by default, if only one task is understandable)
deciding whether or not to attempt to complete the task(s) selected
assessment
(taking stock of what is needed, what one has to work with, and how well one has done)
assessing the characteristics of the test task to determine the desirability and feasibility of successfully completing it and what is needed to complete it
assessing our own knowledge (topical, language) components to see if relevant areas of knowledge are available for successfully completing the test task
assessing the correctness or appropriateness of the response to the test task
planning
(deciding how to use what one has)
selecting elements from the areas of topical knowledge and language knowledge for successfully completing the test task
formulating one or more plans for implementing these elements in a response to the test task
selecting one plan for initial implementation as a response to the test task

Il modello di Bachman è stato usato in modo piuttosto diffuso in letteratura, in particolare nel testing. Esso ha il vantaggio, rispetto ai modelli precedenti, di essere molto più dettagliato nella specificazione delle componenti della competenza linguistica, più preciso nel descrivere il legame che intercorre tra le diverse componenti, più dipendente dalle teorie linguistiche e maggiormente supportato dalla ricerca.

Skehan (1995b; 1998) pur riconoscendo i vantaggi del modello, ne evidenzia però un limite sostanziale: il fatto che esso non sia basato su processi e meccanismi psicolinguistici. Questo rende il modello una valida *checklist*, ma non permette di fare considerazioni rispetto alla natura della performance e del suo legame con la competenza (Skehan, 1998: 164). Rifacendosi a modelli psicolinguistici, Skehan (1995b; 1998) propone di considerare sia come la competenza linguistica viene rappresentata sia i processi legati al suo uso. In particolare, se la relazione tra competenza e performance è dinamica e mediata

dalla competenza strategica, allora è necessario stabilire: (a) se le operazioni attivate dalla competenza strategica sono legate a una rappresentazione mentale astratta (competenza) o piuttosto a processi in tempo reale che influenzano ciò che viene detto (performance); (b) se la consapevolezza della competenza strategica, e le possibili conseguenze di tale consapevolezza sul modo di operare di questa competenza, dipendono dalla competenza o dalla performance; infine (c) che cosa sono esattamente le diverse operazioni attivate dalla competenza strategica e quale processo le coinvolge.

Per capire come la competenza strategica opera può essere utile vedere come la competenza linguistica è rappresentata psicologicamente. In questo senso il lavoro di Skehan condivide una prospettiva di elaborazione dell'informazione dell'uso linguistico (*information processing view*) con altri autori (cfr. per es. Ellis, 1987a; 1987b; Meisel *et al.*, 1981; Van Patten, 1990). Skehan (1985b) riprende la distinzione introdotta da Widdowson (1989) tra analizzabilità e accessibilità. La distinzione di Widdowson è legata alla natura dell'immagazzinamento psicologico (*psychological storage*) del materiale linguistico e alla modalità con cui la comunicazione viene eseguita in tempo reale. L'analizzabilità dipende dal modo con cui l'informazione linguistica è immagazzinata ed è pertanto strettamente collegata alla rappresentazione mentale della conoscenza linguistica stessa. Per l'analizzabilità la regola è di primaria importanza, il sistema pertanto è elegante e compatto e l'unità di riferimento è la parola. L'accessibilità invece ha a che fare con le richieste operative della comunicazione, è legata alla possibilità di accedere alla conoscenza linguistica e interessa l'abilità d'uso, ossia la capacità di mettere in atto la conoscenza. Per l'accessibilità il lessico e le formule sono di primaria importanza, il sistema è facile da usare e soprattutto veloce. Qui l'unità di riferimento è la 'frase lessicalizzata'. Analizzabilità e accessibilità sono dunque i due processi di un *dual-parallel coding system* basato rispettivamente su regole astratte (*rule-based system*) ed esempi (*exemplar-based system*). Poiché le risorse di elaborazione sono limitate, in una situazione comunicativa che richiede un alto grado di *processing*, l'uso formulaico di unità più lunghe delle

parole permette di liberare risorse di attenzione. Il parlante pertanto, cercando di salvaguardare un certo livello di fluenza, tende ad adattare il parlato lungo un continuum di stili di performance. In questo senso la proposta di Skehan (1995b) riprende anche l'idea di interazione tra efficacia comunicativa e accuratezza suggerita da Harold Clahsen, Jürgen Meisel e Manfred Pienemann nel loro *Multidimensional Model* (Clahsen, 1980; Meisel *et al.*, 1981; Pienemann, 1987) secondo cui un apprendente si situa in una determinata posizione del continuum variazionale basandosi sulla sua decisione di distribuire le sue limitate risorse di elaborazione. Si identificano così due estremi del continuum, uno rappresentato da apprendenti che preferiscono evitare deviazioni dalla norma della lingua standard, l'altro da apprendenti che concentrano l'attenzione sull'aspetto comunicativo dell'interazione. La proposta di Skehan si rifà anche allo spettro di stili identificato da Tarone (1988) per l'interlingua, dove l'apprendente si muove tra uno stile più accurato e uno più rilassato e colloquiale. Tuttavia Tarone (1988), come Ellis (1987a; 1987b), al contrario di Skehan, vede questo tipo di variazione sistematica come variazione della competenza:

Planning variability is seen as a feature of the learner's competence, not just of his *performance*. At any stage of development, the learner possesses two or more rules for at least some structures and calls on this differently according to whether the discourse is planned or unplanned. (Ellis, 1987: 14)

Skehan invece interpreta la variazione come parte della performance, qui intesa come esecuzione di un task, quindi la variazione è il riflesso delle priorità di un apprendente rispetto alle sue limitate risorse d'attenzione. Il cambiamento di stile in relazione alle richieste del contesto comunicativo porta così a una ridefinizione del rapporto tra competenza e performance attraverso l'introduzione del costrutto dell'Abilità d'Uso (*Ability for Use*):

Rather than draw upon a generalized and stable underlying competence, the second language performer adjusts to performance conditions by trying to allocate attention in appropriate ways. This enables real-time communication to be more feasible and some degree of fluency achieved. When communicative pressure is not that heavy, when precision is important, or when task demands emphasize form, a syntactic mode assumes greater importance. When pressure is greater and/or when effective communication is paramount, a lexical mode is used more. (Skehan, 1998: 168-69)

L'apprendente dunque attraverso questa competenza di processing è in grado di gestire le diverse richieste della comunicazione rifacendosi a due sistemi di codifica paralleli: la modalità sintattica basata su un sistema di regole (*rule-based system*) e la modalità lessicale basata sugli esempi (*exemplar-based system*). La coesistenza dei due sistemi permette una certa flessibilità nell'uso linguistico. Se la comunicazione può sfruttare sia l'una modalità sia l'altra, o entrambe, allora a volte l'apprendente utilizzerà principalmente la modalità sintattica, e la competenza linguistica sarà il primo fattore di influenza (così come assunto dagli approcci precedenti), a volte invece, quando il tempo reale della comunicazione ha un'influenza significativa, sarà predominante la modalità lessicale, con il risultato che la performance non sarà riflesso diretto della competenza, ma piuttosto di una serie di fattori legati all'elaborazione. Ribaltando così la visione convenzionale che vede la performance basata sulla competenza, ma inevitabilmente influenzata da una serie di fattori che con la competenza non hanno niente a che fare (fatica, complessità del messaggio, stato emotivo, ecc.), Skehan (1995b: 106) con questo approccio duale rafforza il ruolo della competenza di Bachman: "Ability for Use, in other words, is what goes beyond Bachman's (1990) assessment, goal-setting, planning and execution, and is what accounts for the balance between analysability and accessibility as the processing dimension of actual communication". Inoltre, introducendo l'Abilità d'Uso come elemento di mediazione tra le competenze del parlante (cfr. l'analizzabilità e il sistema basato sulle regole) e le richieste del contesto (cfr. l'accessibilità e il sistema basato sugli esempi) focalizza l'attenzione sulle capacità di elaborazione. Sulla base di queste considerazioni Skehan (1998) suggerisce come sia più efficace esaminare direttamente le operazioni attivate dai fattori di processing (che egli identifica nell'*Ability for Use*) piuttosto che le diverse componenti della competenza linguistica.

In questa prospettiva il processo di apprendimento è legato essenzialmente a due operazioni, la ristrutturazione e l'automatizzazione (McLaughlin, 1990). La ristrutturazione è il processo che causa la crescita della complessità sintattica attraverso "the evolution of increasingly abstract

representations of knowledge” (Foster & Skehan, 1996; Schmidt, 1992), mentre l’automatizzazione risulta dal rafforzamento della memoria per formule e routine. Come sostiene Schmidt (1992) si tratta di due processi sostanzialmente diversi ma correlati. La ristrutturazione può essere vista infatti come il risultato dell’analisi di formule sempre più complesse e della loro automatizzazione. E a questo proposito McLaughlin (1990) suggerisce come l’automaticità a livelli inferiori libera risorse di elaborazione per la ristrutturazione a livelli superiori. Analogamente Ellis (1996) ipotizza che la memorizzazione delle formule sia alla base non solo dell’automaticità, ma anche della ristrutturazione. Secondo questa visione l’apprendimento parte dall’esposizione a una forma, procede alla crescente esposizione e memorizzazione di più esempi, e arriva alla generalizzazione, mentre l’ulteriore esposizione porta infine alla memorizzazione delle eccezioni. Così gradualmente si sviluppa una rappresentazione sempre più complessa di una forma che include i diversi esempi, sia regolari sia irregolari, mentre l’accesso a tale forma diviene sempre più automatico.

Skehan (1998) propone di analizzare le produzioni degli apprendenti a partire dalle due dimensioni di processing coinvolte, analizzabilità e accessibilità, e considerando le due modalità, quella sintattica e quella lessicale. Le limitate capacità di memoria richiedono al parlante di scegliere se favorire la forma o il significato. Secondo Skehan, la forma si manifesta nella complessità (se si creano possibilità di *restructuring*, poiché l’apprendente decide di rischiare/sperimentare) o nell’accuratezza (in caso venga data priorità al controllo), mentre il significato è rappresentato dalla fluenza. Sulla base di queste considerazioni Skehan (1998: 270) sostiene che complessità, accuratezza e fluenza possano essere considerati indicatori efficaci nella misurazione della performance in un determinato task.

Ricapitolando, Skehan ridefinisce il rapporto tra competenza e performance introducendo l’Abilità d’Uso, ossia la capacità dell’apprendente di gestire le proprie risorse sulla base delle richieste del contesto. Nel suo modello la variabilità diventa parte integrante della performance e riflette

processi psicolinguistici. In questa prospettiva è dunque possibile analizzare le produzioni degli apprendenti a partire da tre dimensioni: complessità, accuratezza e fluenza. La complessità, ossia il grado di elaborazione della lingua prodotta, viene vista in due modi. In primo luogo, un apprendente può variare nella propensione a usare strutture più complesse e rischiose (Skehan, 2001). In questo caso sono considerate complesse le strutture che si trovano ai limiti del sistema interlinguistico e non ancora completamente automatizzate. In secondo luogo, la complessità si può riferire alla predisposizione dell'apprendente a sperimentare una varietà di strutture. L'accuratezza invece si riferirà a "how well the target language is produced in relation to the rule system of the target language" (Skehan, 1996a: 23). Essa è generata da tre diverse fonti: il grado di accuratezza della rappresentazione stessa, la forza di eventuali rappresentazioni in competizione tra loro, e il grado di automatizzazione della produzione. Durante il processo di apprendimento, infatti, i diversi aspetti della rappresentazione mentale dell'apprendente possono essere più o meno vicini alla lingua target, oltre che differenziarsi per livello di automatizzazione. Gli apprendenti che danno precedenza all'accuratezza esercitano un controllo sugli elementi che hanno internalizzato, adottando un atteggiamento conservativo verso l'uso della L2. La fluenza infine rivela quanto è fluido il parlato di un apprendente, privo di pause o esitazioni inutili. Parte dello sviluppo linguistico, abbiamo visto, è legato chiaramente all'agilità con cui un apprendente può accedere agli elementi linguistici richiesti dall'interazione: la fluenza si ottiene proprio attraverso strategie di processing che permettono all'apprendente di evitare o risolvere problemi velocemente. Viene favorita quando gli apprendenti danno priorità al significato sulla forma così da concludere efficacemente il task.

Rifacendoci a questo approccio in questo studio analizzeremo le produzioni dei soggetti coinvolti in termini di complessità, accuratezza e fluenza. In particolare l'attenzione sarà rivolta alla descrizione della variazione di queste tre dimensioni nel tempo e in diversi task. Poiché il tema della variazione è centrale in questo studio, le sezioni che seguono sono dedicate alla

presentazione della letteratura a proposito in ambito acquisizionale e nel testing, con attenzione specifica agli strumenti di misurazione adottati.

## 2.2. *L'approccio acquisizionale*

Il processo di sviluppo segue percorsi universali comuni a tutti gli apprendenti: l'interlingua di un apprendente in un preciso momento è un sistema coerente e autonomo. Altrettanto riconosciuto il fatto che l'interlingua sia un sistema variabile: nonostante la sistematicità dei percorsi evolutivi, l'interlingua di un apprendente costituisce un sistema instabile, permeabile a nuove forme linguistiche e questa dinamicità si riflette proprio nella variabilità delle produzioni. Descrivere come avviene l'apprendimento della L2 significa dunque spiegare anche il ruolo di tale variabilità in relazione allo sviluppo: la variabilità è 'rumore' da regolarizzare o elemento intrinseco del processo di apprendimento, nonché molla per l'acquisizione? Questa sezione è dedicata al tema della variabilità in prospettiva acquisizionale. Dapprima si discuterà degli approcci utilizzati per descrivere la variazione interlinguistica (§ 2.2.1), poi si affronterà il tema della sua misurazione (§ 2.2.2).

### 2.2.1. La variazione

Sistematicità e variabilità possono apparire come caratteristiche contraddittorie. Eppure nell'ambito dello studio delle L1, sono state individuate modalità per descrivere la variabilità nel parlato dei nativi, interpretabile come risultato dell'applicazione di regole variabili<sup>1</sup> (per es. Labov, 1970; Bayley, 1973) o come grammatiche sovrapposte (*overlapping grammar*) (per es. Bickerton, 1971), dimostrando appunto come sistematicità e variabilità siano in

---

<sup>1</sup> Una regola variabile viene definita nel seguente modo: la regola x viene utilizzata con probabilità X nel contesto sociale Y dato il contesto linguistico Z.



realtà conciliabili. Rifacendosi a questi studi e ai resoconti relativi ai sistemi linguistici dinamici utilizzati nelle scale implicazionali (Bickerton, 1971; De Camp, 1973), a partire dagli anni '80 sono state indagate tre dimensioni di variabilità in ambito acquisizionale: (a) il contesto linguistico in cui una certa struttura viene prodotta, (b) il contesto sociolinguistico con i relativi fattori situazionali, quali il grado di familiarità con l'interlocutore, e (c) il contesto psicolinguistico con i relativi fattori legati alla psicologia della situazione comunicativa, come ad esempio il tempo a disposizione per la pianificazione.

In letteratura possiamo così identificare tre distinti approcci allo studio della variabilità e della sistematicità dell'interlingua: l'*Homogeneous Competence Model* di tradizione chomskiana, l'approccio sociolinguistico e quello psicolinguistico. Secondo il modello chomskiano la variazione è vista come aspetto della performance piuttosto che della competenza, e pertanto esula dalla considerazione del ricercatore, il cui obiettivo è regolarizzare, standardizzare e decontestualizzare le proposizioni parlate e avere accesso così a dati invariabili e dunque rappresentativi della competenza linguistica. Poiché esclude di per sé la variazione, questo approccio non sarà qui considerato. L'approccio sociolinguistico analizza la variazione che ricorre nelle produzioni di un parlante come risultato di differenze nel contesto situazionale, occupandosi dunque non solo di ciò che lo studente sa, ma anche della sua abilità nell'utilizzare tali conoscenze nella comunicazione. Infine l'approccio psicolinguistico intende la variazione come risultato dell'abilità dell'apprendente di elaborare le conoscenze in L2 in diversi contesti d'uso.

Nell'ultimo decennio alcuni ricercatori hanno proposto per lo studio dei processi di acquisizione un ulteriore approccio, detto emergentista (*emergentism*), basato sulla *Dynamic System Theory*, *Chaos/Complexity Theory* e le teorie di *Emergence of Language*. L'apprendimento linguistico viene qui interpretato come un sistema dinamico, ossia un sistema complesso e non lineare in costante mutamento: "each phenomenon is dynamic, complex, nonlinear, chaotic (at times), unpredictable, sensitive to initial conditions, open, self-organizing, feedback sensitive, adaptive, characterized by strange

attractors, which are fractal in shape” (Larsen-Freeman, 1997). Se la sociolinguistica e la psicolinguistica giustificano tale instabilità/variabilità come sistematica e governata da regole, l’approccio emergentista interpreta la sistematicità della variazione come il risultato delle capacità di adattamento del sistema e dunque caratteristica intrinseca del sistema stesso.

Nei prossimi paragrafi si presentano i risultati delle ricerche condotte in ambito acquisizionale. Lo scopo di questo capitolo non è individuare il modello migliore, ma piuttosto evidenziare i diversi aspetti messi in luce nei vari approcci e nelle prospettive teoriche adottate in letteratura. Come suggerisce Zuengler (1989: 66) “it is misguided to search for one comprehensive theory, since one theory will most likely be insufficient in explaining the complexity of performance variation”. In questo studio pertanto ci si riferirà a diverse prospettive teoriche, cercando di cogliere da ognuna strumenti utili per descrivere la variazione del corpus. Dapprima si presenta l’approccio sociolinguistico (§ 2.2.1.1), poi quello psicolinguistico (§ 2.2.1.2), e infine si illustra quello emergentista (§ 2.2.1.3).

#### 2.2.1.1. L’approccio sociolinguistico

I modelli sociolinguistici elaborati in letteratura riguardano principalmente tre paradigmi: quello di William Labov, ripreso da Elaine Tarole e Rod Ellis per la L2, quello dinamico e quello socio-psicologico. Ciascuno di questi paradigmi ha offerto diversi strumenti per l’analisi delle produzioni di apprendenti, contribuendo in modo diverso alla descrizione della variabilità. Il paradigma laboviano ha permesso di esaminare la variazione stilistica sulla base di due costrutti: lo stile discorsivo e la regola variabile; il paradigma dinamico ha posto le basi per l’esplorazione dei rapporti tra forma e funzione e del loro sviluppo nell’interlingua; e infine il paradigma socio-psicologico della *Speech Accommodation Theory* ha contribuito allo studio della variazione legata a

caratteristiche dell'interlocutore e della situazione di interazione. Nei paragrafi che seguono illustro i tre paradigmi in relazione agli studi acquisizionali.

Il primo paradigma che considero è quello di Labov. Nato nell'ambito della L1 ha avuto grande influenza negli studi iniziali sulla variazione in SLA. Due sono gli elementi chiave del paradigma: stile discorsivo e regola variabile. Labov (1970) identifica cinque assiomi legati allo stile discorsivo: (a) non c'è un unico stile discorsivo; (b) gli stili possono essere ordinati rispetto al grado di attenzione verso il parlato; (c) lo stile colloquiale è quello in cui si presta meno attenzione al *monitoring*; (d) lo stile colloquiale non può essere osservato in contesti formali; (e) l'unico modo per ottenere 'buoni dati' è l'osservazione sistematica. Le regole variabili registrano invece tre diversi aspetti: (a) il contesto linguistico, (b) il contesto sociale in cui una forma linguistica viene utilizzata, e (c) la probabilità di applicazione di una regola. Tarone (1985) e Ellis (1985a; 1985b) riprendendo il lavoro condotto da Labov propongono due modelli teorici per l'interpretazione della variabilità dell'interlingua.

Tarone (1983) con il *Continuous Competence Model* sostiene che l'interlingua di un apprendente costituisce un continuum di stili discorsivi che differiscono tra loro per grado di variabilità e coerenza interna. Tarone (1988) presenta 75 studi che mirano a indagare la variabilità dell'interlingua. Buona parte di queste ricerche si riferisce alle differenze dovute al contesto comunicativo, ossia come cambiamenti nelle caratteristiche dei task e nelle procedure di elicitazione possano portare a cambiamenti nel livello di accuratezza d'uso di determinate strutture a tutti i livelli linguistici: fonologico, morfologico, sintattico o lessicale. A livello fonologico, ad esempio, Dickerson (1975) mostra come giapponesi apprendenti di inglese varino in accuratezza la pronuncia di /r/ con il variare del task: lettura di una lista di parole, lettura di un dialogo e conversazione libera, raggiungendo un livello di accuratezza inferiore nella conversazione libera e superiore nella lettura di una lista di parole. A livello morfologico, Tarone (1985) ad esempio mostra come apprendenti giapponesi e arabi di inglese variano il grado di accuratezza nell'uso del morfema -s del plurale con il variare del task (per es., test

grammaticale a scelta multipla, intervista e racconto di una storia) ottenendo un maggior livello di accuratezza nel test grammaticale. Opposto il pattern di variazione per l'articolo: in questo caso il maggior livello di accuratezza si riscontra nel task racconto di una storia. A livello sintattico Liu (1991) e Tarone & Liu (1995), in uno studio longitudinale sullo sviluppo interlinguistico di un bambino cinese di 6 anni in tre contesti differenti per grado di familiarità con l'interlocutore, mostrano come lo sviluppo delle forme interrogative sia diverso nei tre contesti situazionali considerati. Ogni nuova forma interrogativa si sviluppa prima nell'interazione con un adulto familiare in contesto di gioco a casa, poi nell'interazione con i compagni in classe, e solo alla fine nell'interazione con l'insegnante.

Nel modello elaborato da Elaine Tarone la scelta dello stile dipende principalmente dalla quantità di attenzione diretta alla forma linguistica rispetto al contenuto: maggior attenzione alla forma porterebbe a una riduzione della variabilità. L'acquisizione pertanto procede in due modi diversi. Da un lato, le nuove forme interlinguistiche sono inizialmente interiorizzate quando il parlante partecipa a conversazioni spontanee, poi dallo stile colloquiale si propagano negli stili più accurati. Dall'altro, le forme che vengono inizialmente prodotte negli stili più accurati (perché apprese attraverso l'insegnamento) si propagano poi nello stile colloquiale. La variazione viene così strettamente collegata all'acquisizione.

Ellis (1985a; 1985b) propone il *Variable Competence Model*. Questo modello identifica due tipi di variabilità dell'interlingua: la variabilità sistematica, che include quella individuale e del contesto, e la variabilità non sistematica, che include quella libera e quella dovuta alla performance. Nell'un caso l'interlingua è costituita da regole in competizione tra loro e strettamente collegate alla situazione o a fattori contestuali; nell'altro l'alternarsi di regole diverse è invece casuale, e la grammatica dell'apprendente è indipendente da contesto e situazione. L'acquisizione, secondo Ellis, avverrebbe in due tappe: prima si assimilano nuove forme, poi si ricostruisce la relazione forma-funzione. La variazione libera indica dunque che l'apprendente ha individuato

una nuova forma, ma non l'ha ancora messa in relazione con la sua funzione specifica. Per Ellis (1999) la variazione libera è dunque la molla che fa scattare l'acquisizione.

Nei due modelli di Elaine Tarone e di Rod Ellis la competenza viene definita come variabile. Tale posizione non è stata esente da critiche. Gregg (1990) ad esempio, fedele alla sua formazione generativista, pur riconoscendo la variabilità delle performance come fatto incontrovertibile, pone alcune questioni metodologiche, e mette in dubbio l'utilità stessa di includere la variabilità in una teoria dell'apprendimento della L2. Ai due autori rimprovera una certa vaghezza nella definizione del costrutto competenza variabile:

in principle they claim that linguistic knowledge is itself variable for any given learner, but in practice they avoid any concrete discussion of what that knowledge consists of, leaving the 'variable competence' construct up in the air and confining themselves to (often extensive, detailed, and interesting) discussion of output data. (Gregg, 1990: 378)

Secondo Gregg il concetto di competenza variabile rischia infatti di confondere i confini tra conoscenza linguistica e produzione, diventando poco coerente dal punto di vista teorico: se la variazione è parte della performance, è legittimo parlare di una teoria della variazione dell'interlingua all'interno di una teoria della performance, ma non lo è farlo entro una teoria che mira a descrivere lo sviluppo della competenza.

Il secondo paradigma sociolinguistico è quello dinamico. In questo tipo di ricerche l'attenzione è sullo studio dei rapporti forma-funzione. La variazione interlinguistica non è qui messa in relazione con un'idea di competenza variabile. Preston (1993: 170) ad esempio suggerisce come "there is no inconsistency in believing that variation arises from alternatives (not 'variation') in competence and that the probabilistically determined realisation of those alternatives is due to a variety of contributing factors, not necessarily to a built-in variability in linguistic competence itself". Nemmeno l'idea di variazione libera introdotta da Rod Ellis è condivisa in questo approccio: la variazione non sistematica è considerata solo come apparente, probabilmente frutto di metodi di ricerca o corpora di dati inadeguati (Preston, 1996: 25).

In questo paradigma la variazione nel parlato viene indagata utilizzando le stesse tecniche impiegate in sociolinguistica per studiare la variazione tra parlanti nativi (VARBRUL e analisi multivariata), e mostrando come i vincoli che determinano la variazione siano ordinabili sulla base del peso che hanno nel definire la variazione orizzontale (Bayley & Preston, 1996; Bayley & Langman, 2004). Vediamo alcuni esempi. Wolfram (1985) studiando apprendenti vietnamiti d'inglese mostra come la marca del tempo passato sia sistematicamente influenzata da fattori quali la salienza, definita come differenza tra forma base e forma passata del verbo, e, nel caso dei verbi irregolari, da aspetti del contesto fonologico. Young (1991) ha utilizzato l'analisi variazionista (*variable rule analysis*) per studiare la marca del plurale da parte di cinesi adulti apprendenti di inglese. Il suo studio dimostra come l'uso del plurale sia soggetto a una serie di vincoli, tra cui il livello di proficiency, la ridondanza della marca di plurale, la funzione sintattica del sintagma nominale, le caratteristiche dei segmenti che seguono e precedono. Bayley (1991; 1994), Langman & Bayley (2002) e Howard (2004) utilizzando l'analisi multivariata mostrano come l'utilizzo della marca del plurale sia influenzato da una serie di fattori sia linguistici che sociali: il passato viene impiegato preferibilmente con verbi perfettivi, mentre la forma non flessa con verbi imperfettivi; inoltre tale marca può variare in relazione alla regolarità/irregolarità del verbo in questione, la scelta del passato è influenzata dalla salienza e dal gruppo sociale di appartenenza dell'apprendente. Anche in questi studi la variazione sincronica è vista come precursore del cambiamento, come ad esempio illustra Bayley (1973) nel *Wave Model*: inizialmente una regola ha possibilità di applicazione zero in tutti i contesti, poi nel tempo la probabilità di applicazione cresce contesto per contesto.

A partire dagli anni '90 si sviluppa entro questo paradigma un filone di ricerca che analizza le produzioni di parlanti nativi come controllo in relazione al quale valutare l'acquisizione della medesima variazione sociolinguistica da parte degli apprendenti. Così per esempio Adamson & Regan (1991) studiano l'acquisizione da parte di immigrati vietnamiti e cambogiani del pattern

variabile nativo rispetto all'uso del morfema *-ing* (per es. *working* vs *workin'*) e riscontrano un effetto della variabile genere sulla scelta effettuata dal parlante, nel senso che le donne usano [*-in*] meno frequentemente degli uomini, soprattutto nel parlato controllato. Bayley (1996) invece studia l'acquisizione da parte di cinesi adulti del pattern nativo inglese della riduzione del nesso consonantico e constata come il livello di proficiency e la quantità di contatti con i parlanti nativi correlino con la variazione. Major (2004) studia l'acquisizione di forme fonologiche native da parte di giapponesi e spagnoli apprendenti di inglese, e trova differenze interessanti in relazione al genere degli apprendenti e allo stile che usano, suggerendo che la variazione preceda quella di stile, nel senso che le donne usano meno forme colloquiali degli uomini. Regan (1995; 1996; 1997; 2004) studia l'acquisizione da parte di alcuni studenti irlandesi del pattern nativo di cancellazione della particella *ne* nella negazione francese, e mostra come, dopo un anno di residenza in Francia, gli apprendenti comincino ad avvicinarsi all'uso nativo (vedi anche Dewaele, 1992; Dewaele & Regan, 2002). Sankoff *et al.* (1997) studiano l'uso dei segnali discorsivi francesi da parte di parlanti non nativi residenti a Montreal, e conclude come l'uso appropriato dei segnali discorsivi sia indice di un buon grado di integrazione con la comunità francofona. In una serie di studi riferiti al francese canadese Mougeon & Rehner (2001), Rehner & Mougeon (1999) e Rehner *et al.* (2003) analizzano 13 variabili distribuite lungo il continuum sociolinguistico formale-informale, ed evidenziano come alcuni apprendenti inseriti in programma di immersione non usino le forme più colloquiali, ma preferiscano forme intermedie. Le ricerche qui presentate concordano nel dimostrare che la variazione nell'interlingua, così come nella lingua parlata dai nativi stessi, è sistematica e regolata da diversi vincoli. Di conseguenza apprendere una L2 significa apprenderne anche i pattern di variabilità (Rehner & Mougeon, 1999).

Un terzo approccio alla variazione è rappresentato dai modelli di tipo socio-psicologico, come ad esempio la *Speech Accommodation Theory* di Giles (1977). La variazione d'uso linguistico qui è messa in relazione con lo stato

sociale, istituzionale ed etnico del gruppo di appartenenza del parlante. In base al tipo di aggiustamenti che il parlante fa durante lo scambio comunicativo vengono identificate tre dimensioni di variabilità: (a) la convergenza, nel caso in cui il parlante adatti il suo discorso per renderlo più simile a quello dell'interlocutore; (b) la divergenza, quando il parlante differenzia il suo eloquio rispetto a quello dell'interlocutore; e (c) la costanza (*speech maintenance*), nel caso in cui il parlante non modifica il suo stile. Uno dei punti forti di questo modello riguarda il riconoscimento del ruolo centrale dell'interlocutore nel determinare la variazione della produzione. L'accomodamento è identificabile a tutti i livelli linguistici, e può dunque, in contesti bilingui, riguardare tanto la scelta stessa del codice, quanto la velocità, il tono, la pronuncia, le scelte lessicali o grammaticali, nonché elementi del discorso, quali la lunghezza dei turni o la scelta dell'argomento di conversazione. Usata in linguistica acquisizionale da ricercatori quali Beebe (1981; 1987), Beebe & Zuengler (1983) e Zuengler (1989), la *Speech Accommodation Theory* ha contribuito sostanzialmente allo studio della variazione linguistica legata al ruolo dell'interlocutore. Beebe (1985) in particolare riconosce l'esistenza di un'ulteriore dimensione specifica dell'interlingua, che definisce *extravergence*, ossia quella variazione che non rappresenta né convergenza, né divergenza, ma piuttosto è legata all'apprendimento, alla comunicazione e alle strategie di performance o alla variazione sistematica connessa al contesto linguistico tipiche dell'interlingua.

Ricapitolando, i modelli sociolinguistici qui presentati hanno offerto ai ricercatori di SLA una serie di strumenti teorici e metodologici utili per lo studio dell'interlingua. L'idea di sistematicità della variazione e l'identificazione dei fattori che la determinano sono i due elementi di maggior interesse ai fini di questo studio. Non si farà riferimento invece alla metodologia di analisi variazionista. L'approccio sociolinguistico infatti essendo legato alla visione della variazione come fenomeno di gruppo, piuttosto che individuale, può avere limiti quando applicato allo studio della



variazione individuale, obiettivo degli studi SLA. Più utile ai nostri fini l'approccio psicolinguistico, che illustro di seguito.

#### 2.2.1.2. L'approccio psicolinguistico

Negli studi di stampo sociolinguistico o socio-psicologico, come abbiamo visto in § 2.2.1.1, la variabilità interlinguistica è intesa come variazione sistematica dell'accuratezza fonologica o morfo-sintattica con il mutare di caratteristiche situazionali specifiche quali il tipo di attività svolta, il grado di attenzione rivolto alla forma piuttosto che al significato o il grado di familiarità con l'interlocutore. Vediamo ora gli studi che si sono occupati di modelli di variazione secondo un approccio psicolinguistico analizzando la variazione in termini di processi mentali richiesti nella produzione del parlato.

Alcuni autori, come Meisel *et al.* (1981) e Pienemann (1988; 1992) elaborando un modello multidimensionale spiegano la variazione dell'interlingua in termini di utilizzo di diverse componenti del sistema cognitivo di produzione linguistica nel realizzare diversi task. Pienemann raffina tale modello proponendo la Teoria della Processabilità (d'ora in poi TP). La TP come approccio permette di fare predizioni sulle traiettorie del percorso di sviluppo nell'acquisizione di qualsiasi L2. La nozione di traiettoria di sviluppo (*developmental trajectory*) implica la presenza sia di una dimensione sistematica e universale (*staged development*) sia di una dimensione di variazione che giustifica le differenze nei percorsi di sviluppo a livello individuale. Secondo questo approccio è possibile individuare diverse traiettorie di sviluppo basate sullo stesso stadio: durante il processo di acquisizione, l'apprendente accumula le regole grammaticali e le loro varianti, percorrendo così una traiettoria di sviluppo individuale, aderendo comunque a una sequenzialità universale. In questo modo la TP definisce uno spazio bidimensionale per lo sviluppo delle ipotesi (per es. *Hypothesis Space*), entrambe le dimensioni di questo spazio sono regolate dalla gerarchia della

processabilità (Pienemann, 2008: 10-1). In questo senso per Pienemann “differences in the percentage of rule application are understood as accuracy differences which do not alter the basic nature of the maturity of the IL system” (1992: 7). Poiché l’acquisizione è definita in termini di emersione, la variabilità è ricondotta entro l’universalità del percorso (Pienemann, 1998). Pienemann suggerisce applicazioni pratiche del suo modello multidimensionale, includendo il testing e suggerendo come la TP possa fornire un’utile base teorica a chi si occupa di valutazione (cfr. Brindley, 1998). La posizione espressa da Pienemann non è esente da critiche. Le potenzialità della teoria sembrano infatti scontrarsi con alcune difficoltà legate alla sua operativizzazione, con particolare riferimento a due aspetti: la definizione operativa del costrutto emersione (cfr. Pallotti, 2007) e la selezione di task efficaci e replicabili. Se una determinata struttura emerge in momenti diversi con il variare del contesto situazionale (cfr. Liu, 1991; Tarone & Liu, 1995), sarà essenziale prestare attenzione al contesto di elicitazione, sia per quel che riguarda la selezione di task rappresentativi, sia rispetto alla necessità di mantenerli confrontabili nei diversi studi. Poiché la TP riconduce la variabilità individuale entro un percorso universale attraverso il costrutto emersione, questo approccio non sarà qui ulteriormente considerato.

Un ulteriore gruppo di studi indaga la variazione in termini di processi mentali. Qui si inseriscono da un lato le ricerche sui modelli di pianificazione (per es. Crookes, 1989; Ellis, 1987a; 1987b; Yuan & Ellis, 2003; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Wendel, 1997), dall’altro i modelli di *processing* dell’informazione cognitiva (*cognitive information processing*) dedicati all’analisi dell’uso delle risorse di attenzione durante il completamento di un task, tra i quali in particolare il *Limited Attentional Capacity Model* (Skehan, 1998; 2001; 2003; Skehan & Foster, 1999; 2001) e la *Cognition Hypothesis* (Robinson, 2001a; 2001b; 2003a; 2005a; 2007b). Poiché questo studio si inserisce entro questo quadro teorico nell’indagare la variazione, i paragrafi che seguono descrivono i modelli in dettaglio.

Il modello per la produzione del parlato proposto da Levelt (1989)

suggerisce una serie di possibili cause psicolinguistiche di variazione. Nella produzione del parlato, in un primo tempo nel concettualizzatore si effettuano decisioni sulla varietà linguistica da utilizzare sulla base di fattori situazionali e intenzioni comunicative, poi nel formulatore il messaggio pre-verbale viene convertito in discorso pianificato tramite la selezione del lessico e della sua messa in grammatica, e infine nell'articolatore viene tramutato in onda sonora. A questo sistema di produzione se ne affianca poi uno di comprensione, che durante la produzione fornisce al parlante il *feedback* necessario per mettere a punto eventuali adattamenti di diverso tipo.

Le ricerche condotte in questo ambito si occupano in modo particolare della nozione di attenzione, intesa come meccanismo psicolinguistico attraverso cui elementi sociali, stilistici o cognitivi influiscono sulla performance. Questi studi si rifanno in modo più o meno esplicito all'*information processing theory*, che sostiene come i parlanti abbiano limitate capacità di processing (Miyake & Shah, 1999; Baddeley, 2007) e di conseguenza non siano in grado di controllare tutti gli aspetti di un task (Anderson, 1995; Newell & Simon, 1972). Gli apprendenti, in particolare quelli ai livelli più bassi, possono avere difficoltà a prestare attenzione contemporaneamente a forma e significato, e devono decidere come distribuire le proprie risorse, favorendo un aspetto piuttosto dell'altro (Anderson, 1995; Skehan, 1996a; VanPatten, 1990). Abbiamo già introdotto a proposito in § 2.1.2 i lavori di Skehan, secondo cui la variazione della performance è caratterizzata da tre dimensioni: la complessità, ossia l'elaborazione della lingua prodotta, l'accuratezza, ossia il grado di conformità delle produzioni alla norma linguistica della lingua target, e la fluenza, ossia la capacità di produrre un discorso in tempo reale senza eccessive pause o esitazioni (Skehan, 1996a: 22). Secondo Skehan (1998), queste tre dimensioni della performance sono rilevanti in quanto influenzate dal tipo di *processing* attivato dall'apprendente. Rifacendosi a due sistemi di codifica paralleli, modalità sintattica (basata su *exemplar-based system*) e modalità lessicale (basata su *rule-based system*), il parlante gestisce le proprie risorse per rispondere alle richieste della situazione

comunicativa. In determinate condizioni, l'apprendente deciderà di favorire il significato affidandosi al sistema basato sugli esempi, con un conseguente aumento della fluenza, in altre condizioni cercherà di favorire la forma appoggiandosi maggiormente al sistema basato sulle regole, ottenendo così una maggiore complessità e/o accuratezza (cfr. § 2.1.2).

Un primo gruppo piuttosto cospicuo di studi si è occupato dell'impatto della pianificazione sulla produzione linguistica (cfr. Ellis, 2005; Ortega, 1999; Bygate *et al.*, 2001; per una rassegna dei risultati; cfr. tabella 1 in Appendice). Abbiamo visto che poiché hanno risorse limitate i parlanti possono avere difficoltà a prestare attenzione contemporaneamente a forma e significato, quando però hanno la possibilità di pianificare un task sono in grado di compensare le proprie limitate capacità di processing e ottenere una produzione qualitativamente migliore (Skehan, 1996). I risultati di queste ricerche dimostrano in generale un effetto positivo della pianificazione con un aumento della fluenza nelle produzioni precedute da maggiore pianificazione (per es. Crookes, 1989; Foster & Skehan, 1996; Gass *et al.*, 1999; Tavokoli & Skehan, 2005). In modo analogo un effetto positivo della pianificazione è stato osservato anche rispetto alla complessità (per es. Crookes, 1989; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Yuan & Ellis, 2003; Skehan & Foster, 2005), mentre risultati meno chiari sono stati ottenuti rispetto all'accuratezza (per es. Ellis, 1987; Crookes, 1989; Foster & Skehan, 1996; 1997; Ortega, 1999; Wigglesworth, 1997; Bygate, 2001).

Un punto di interesse comune a queste ricerche sull'effetto della pianificazione riguarda l'attenzione rivolta agli effetti di compensazione (*trade-off*) tra complessità, accuratezza e fluenza, anche se in questo senso i risultati sono in parte contrastanti. Skehan & Foster (1997) ad esempio osservano un effetto compensatorio tra complessità e accuratezza, nel senso che quando cresce la prima la seconda cala. Nello studio di Mehnert (1998) una breve pianificazione sembra favorire l'accuratezza, e una pianificazione più lunga sia l'accuratezza sia la complessità. Infine nelle ricerche condotte da Yuan & Ellis (2003) si osserva che mentre la fluenza aumenta con la possibilità

di pianificare la produzione linguistica prima del task, l'accuratezza aumenta quando questa possibilità esiste *on-line*, ossia durante la produzione. Se in letteratura è ormai assodato che la pianificazione prima di svolgere il task aiuta gli apprendenti a produrre un linguaggio più fluente e complesso, è meno chiaro se favorisca anche l'accuratezza. Alcuni studi suggeriscono che in questo caso sia più efficace la pianificazione *on-line*.

Una seconda area che ha stimolato diverse ricerche riguarda la ripetizione del task (cfr. Tabella 2 in Appendice). Gass *et al.* (1999) osservano come la ripetizione del task abbia effetti positivi sulla qualità della performance anche se non durevoli nel tempo. Analogamente Bygate (1996; 1999; 2001) in una serie di studi riporta come la ripetizione favorisca la complessità e la fluenza, mentre i risultati per l'accuratezza non sono sempre statisticamente significativi. Lynch & McLean (2001) lavorando con studenti di corsi di medicina riscontrano come nella ripetizione del task 'presentazione di un poster' gli apprendenti più avanzati riescano a migliorare le loro presentazioni e cogliere spunti dai contributi degli interlocutori. Nemeth & Kormos (2001) usando task argomentativi osservano come la ripetizione porti a una maggior attenzione verso il contenuto e l'uso di forme di supporto.

Un terzo gruppo di studi si rifà a un approccio cognitivo di elaborazione dell'informazione, dedicandosi all'analisi dell'uso delle risorse di attenzione durante il completamento di un task, della relazione tra caratteristiche del task e caratteristiche della performance, e dell'impatto delle diverse condizioni in cui viene realizzato il task (cfr. Tabella 3 in Appendice). In questo filone esistono due modelli: il *Limited Attentional Capacity Model* (Skehan, 1998; 2001; 2003; Skehan & Foster, 1999; 2001) e la *Cognition Hypothesis/Triadic Componential Framework* (Robinson, 2001a; 2001b; 2003; 2005; 2007). Entrambi descrivono come la performance viene influenzata dalle caratteristiche dei task e dalle condizioni in cui sono svolti, e come le risorse di attenzione hanno un ruolo chiave durante lo svolgimento di un task, mentre raggiungono conclusioni diverse sugli effetti di alcune caratteristiche dei task sulla performance. La tabella 1 mette a confronto i due modelli.

**Tabella 1. I modelli cognitivi di elaborazione dell'informazione**

Limited Capacity Model, da Skehan, 1998		
code complexity	cognitive complexity	communicative stress
linguistic complexity and variety vocabulary load and variety redundancy and density	<i>cognitive familiarity</i> familiarity of topic and its predictability familiarity of discourse genre familiarity of task  <i>cognitive complexity</i> information organisation amount of computation clarity and sufficiency of information given information type	time limits speed of presentation number of participants length of text used type of response opportunities to control interaction
Triadic component framework, da Robinson, 2005		
task complexity	task conditions	task difficulty
(a) <i>resource-directing variables</i> ± few elements ± here and now ± no reasoning demands	(a) <i>participation variables</i> , e.g. open solution one way/two way convergent/divergent	<i>ability variables</i> , e.g. working memory intelligence aptitude
(b) <i>resource-dispersing</i> ± planning time ± single task ± prior knowledge	(b) <i>participant variables</i> , e.g. same/different gender familiar/unfamiliar power/solidarity	<i>affective variables</i> motivation anxiety confidence
Triadic component framework, da Robinson, 2007		
task complexity	task conditions	task difficulty
(a) <i>resource-directing variables making cognitive/conceptual demands</i> ± here and now ± few elements ± spatial reasoning ± causal reasoning ± intentional reasoning ± perspective-taking	(a) <i>participation variables making interactional demands</i> ± open solution ± one way flow ± convergent solution ± few participants ± few contributions needed ± negotiation not needed	(a) <i>ability variables and task relevant resource differentials</i> h/l working memory h/l reasoning h/l task switching h/l aptitude h/l field independence h/l mind reading
(b) <i>resource-dispersing variables making performative/procedural demands</i> ± planning time ± prior knowledge ± single task ± task structure ± few steps ± independency of steps	(b) <i>participant variables making interactant demands</i> ± same proficiency ± same gender ± familiar ± shared content knowledge ± equal status and role ± shared cultural knowledge	(b) <i>affective variables and task relevant state-trait differentials</i> h/l openness h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to communicate h/l self-efficacy

Secondo il *Limited Attentional Capacity Model* (Skehan, 1998; 2001; 2003; Skehan & Foster, 1999; 2001) le risorse di attenzione attivate dalle *resource-directing variables* sono limitate, e di conseguenza l'aumento della complessità di un task può ridurre la quantità di attenzione disponibile. Quando i limiti di carico di attenzione vengono raggiunti, gli apprendenti dovranno decidere se dare priorità al significato o alla forma. In altre parole, con l'aumento della complessità cognitiva del task gli apprendenti saranno costretti ad adottare una di queste due strategie: prestare attenzione al significato prima che alla forma, con una conseguente riduzione della complessità e dell'accuratezza della loro produzione a favore della fluenza, oppure prestare attenzione alla forma o alla sperimentazione di nuove strutture, con un aumento di complessità o accuratezza, a scapito della fluenza. Robinson (Robinson, 2001a; 2001b; 2003; 2005; 2007) con la sua *Cognition Hypothesis* sostiene una posizione in parte contrastante. Innanzitutto le risorse di attenzione attivate dalle *resource-directing variables* non sono limitate, come sostenuto da Skehan, anzi gli apprendenti possono sfruttare tipi di attenzione che non sono in competizione tra di loro (*non-competing attentional pools*). Di conseguenza, l'aumento della complessità cognitiva del task può portare a una maggiore complessità e accuratezza della produzione. In altre parole, l'aumento della richiesta cognitiva di un task può spingere le risorse di attenzione di un apprendente verso la forma, così che l'input viene processato in modo più profondo ed elaborato. Entrambi hanno stimolato diverse ricerche che indagano gli effetti delle caratteristiche e condizioni dei task sulla performance. Vediamo ora i risultati ottenuti (cfr. tabella 1 in Appendice per una ricognizione degli studi relativi al parlato).

Peter Skehan e Pauline Foster in una serie di studi dimostrano come alcune caratteristiche dei task influenzino la performance: un task strutturato e familiare favorisce una maggiore fluenza e accuratezza, ma non necessariamente la complessità (Foster & Skehan, 1996; Skehan & Foster, 1997; 1999); un task che richiede una giustificazione come risultato produce

lingua più complessa (Skehan & Foster, 1997); i task interattivi portano a produzioni più accurate e complesse, i task monologici favoriscono invece la fluenza (Foster & Skehan, 1996; 1999; Skehan & Foster, 1997; 1999). L'aumento della complessità del task lungo la dimensione *resource-directing* genera una maggior quantità di interazione e negoziazione, quindi più *uptake* e incorporazioni dell'input, favorendo così una produzione più complessa, ma meno fluente per i task più complessi (Robinson, 2000; 2001). Le ricerche che hanno indagato la dimensione *Here-and-Now/There-and-Then* osservano gli effetti positivi della complessità del task sull'accuratezza (Iwashita *et al.*, 2001; Rahimpour, 1997; 1999; Gilabert, 2007; Ishikawa, 2007). In una serie di studi sulla produzione scritta (Kuiken *et al.*, 2005; Kuiken & Vedder, 2004a; 2004b; 2007a; 2007b; 2007c; 2008) in cui si osservano gli effetti del fattore  $\pm$  elementi l'aumento della complessità del task genera produzioni più accurate e una maggior varietà lessicale, mentre non si riscontrano effetti rispetto alla complessità sintattica. Michel *et al.* (2007) indagando l'effetto dei fattori  $\pm$  elementi e  $\pm$  interattivo osservano come il task dialogico favorisca accuratezza e complessità lessicale, mentre il task monologico porta a un maggior uso di riparazioni e a un aumento di complessità sintattica e lessicale. In generale c'è poco supporto rispetto alla possibilità che un task complesso generi una sintassi più complessa, maggior supporto rispetto a un aumento dell'accuratezza. Robinson (2005) basandosi sullo studio di Niwa (2000) suggerisce come le differenze individuali potrebbero mascherare gli effetti della variazione del task.

Nonostante la vivacità delle ricerche in questa direzione è probabilmente prematuro cercare di appoggiare definitivamente uno dei due modelli. La difficoltà di trovare una soluzione è legata a diversi fattori. Le ricerche ad oggi condotte hanno utilizzato task e condizioni molto diversi, per cui i confronti tra ricerche non sono scontati. Poiché manca ancora in letteratura un accordo su misure standardizzate, i vari studi utilizzano diverse misure per operationalizzare complessità, accuratezza e fluenza (CAF), rendendo il confronto tra risultati non sempre affidabile. Nonostante questi limiti, gli studi



condotti con misure che rilevano queste tre dimensioni della produzione sembrano promettenti e utili, questo studio cerca pertanto di offrire un contributo in questo senso: misurare CAF nelle produzioni elicitate per mezzo di task è utile perché può aiutare a definire un modello di proficiency nella seconda lingua con dirette applicazioni in campo didattico (cfr. Ellis, 2003) e nel testing (cfr. Iwashita *et al.*, 2001).

#### 2.2.1.3. L'approccio emergentista

L'approccio emergentista riconsidera in modo sostanziale il ruolo della variazione rispetto agli approcci visti in precedenza. Se questi ultimi vedevano la comunicazione come un sistema caratterizzato da una sequenza lineare di eventi, che permette lo scambio di informazioni tra mittente/parlante e ricevente con un meccanismo simile a quello di una macchina per fax, dove i due partner della conversazione si alternano nell'emettere e processare il messaggio, l'approccio emergentista si rifà ai sistemi dinamici e descrive la comunicazione con la metafora della danza. A partire da semplici procedure realizzate in coppie coordinate, come per i passi nella danza, pattern complessi emergono nel tempo, con pattern ancora più complessi che si realizzano quando coppie di ballerini interagiscono con altre coppie (cfr. de Bot *et al.*, 2007: 9). Se l'approccio tradizionale è legato a termini quali segnale e risposta, inviare e ricevere, codificare e decodificare o a comportamenti basati su regole, l'approccio emergentista usa termini quali *engagement* e *disengagement*, sincronia e asincronia, *breakdown* e *repair* nell'interazione per descrivere i processi di comunicazione.

Lo studio dei sistemi dinamici, ossia sistemi complessi e non lineari in costante mutamento, è nato nell'ambito della matematica e delle scienze con la *Dynamic System Theory* (d'ora in poi DST) e la *Chaos/Complexity Theory* (d'ora in poi C/CT). Negli ultimi anni però ha avuto diverse applicazioni anche nelle discipline umanistiche (cfr. van Geert, 1991; Thelen & Smith, 1994),

inclusa la linguistica acquisizionale: con la C/CT proposta da Larsen-Freeman (1997; per una presentazione della letteratura, Larsen-Freeman & Cameron, 2008), la DST illustrata da de Bot *et al.* (2007; cfr. Herdina & Jessner, 2002), e le teorie emergentiste *Emergence of Language* introdotte da Ellis (1996; 1998) e MacWhinney (1998; 1999; 2001).

Sulla base di possibili somiglianze tra sistemi dinamici e apprendimento linguistico, Diane Larsen-Freeman con l'articolo *Chaos/Complexity science and second language acquisition* del 1997 ha proposto l'applicazione dell'approccio emergentista ai processi di apprendimento di una seconda lingua, evidenziando come questo approccio possa essere un utile strumento per descrivere in modo più realistico il processo di apprendimento, includendo aspetti sia cognitivi che sociali. Dal 1997 ad oggi diversi studi sono stati condotti in questa direzione e alcuni autori hanno proposto la C/CT come possibile teoria generale dello sviluppo linguistico (cfr. Larsen-Freeman, 2002; *Applied Linguistics, Special Issue 27*, 2006; *Bilingualism: Language and Cognition, Special Issue 10*, 2007; Larsen-Freeman & Cameron, 2008; per una presentazione della letteratura in proposito). In questa sezione presento le caratteristiche principali dell'approccio emergentista focalizzando l'attenzione in particolare su variazione e andamento dei percorsi di sviluppo.

Un sistema dinamico è caratterizzato da alcune proprietà distintive: (a) completa interconnessione dei sotto-sistemi, tutte le variabili sono legate l'una all'altra, pertanto il cambiamento di una di esse ha un impatto su tutte le altre. Un esempio il cosiddetto *butterfly effect*, dipendenza dallo stato iniziale, lo sviluppo dei sistemi dinamici è strettamente legato allo stato iniziale del sistema, pertanto piccole differenze iniziali possono avere conseguenze molto evidenti sul lungo periodo. Si generano così traiettorie di sviluppo non facilmente prevedibili. L'interazione tra variabili porta a cambiamenti che non possono essere anticipati con gli strumenti analitici classici; (b) l'emersione di stati di attrazione (*attractor states*) o stati di repulsione (*repeller states*), poiché i sistemi dinamici sono sempre parte di altri sistemi più ampi, i diversi sotto-sistemi si regolano su determinati stati, per definizione temporanei e non fissi,

detti appunto stati di attrazione, preferiti ma non forzatamente prevedibili, o stati di repulsione, non preferiti; infine (c) evoluzione costante (*constantly changing*), i sistemi dinamici sono in costante mutamento, grazie all'interazione con il contesto e attraverso una costante riorganizzazione interna: tale condizione rende i sistemi dinamici per definizione variabili.

Riprendendo questo approccio, van Geert (1991) propone un modello di sviluppo della L1 basato sulla DST. Perché lo sviluppo avvenga è necessario che una serie di condizioni siano rispettate. In primo luogo deve esserci qualcosa che può crescere, in questo senso van Geert parla di *minimal structural growth condition*; in secondo luogo, dovranno esserci risorse che sostengono il processo di crescita. In questo caso van Geert distingue tra risorse interne, legate all'individuo, quali capacità o tempo per apprendere o elementi quali conoscenze concettuali e motivazione, e risorse esterne, come ad esempio contesto di apprendimento o eventuali interventi didattici. Queste risorse hanno due caratteristiche fondamentali: sono limitate e sono interconnesse, da un lato dovranno essere posizionate su un valore minimo perché avvenga l'apprendimento, dall'altro, quando necessario, il sistema ha a disposizione meccanismi di compensazione che coinvolgono diversi tipi di risorse contemporaneamente. Van Geert (1994) chiama *connected growers* proprio quei sottosistemi che sostengono la reciproca crescita. È il caso ad esempio della relazione tra sviluppo lessicale e sviluppo della capacità di comprensione: in genere conoscere più parole rende la comprensione del parlato più semplice. Poiché le risorse sono limitate, anche la crescita per definizione è limitata. Van Geert parla in questo caso di *carrying capacity*, riferendosi appunto allo stato di conoscenza che può essere raggiunto da un determinato bambino nello sviluppare strutture tra loro collegate. È il caso ad esempio dell'emersione delle frasi complesse: questa può determinare un rallentamento dello sviluppo lessicale. Se prima infatti tutte le risorse erano impiegate nello sviluppo del lessico, in questa seconda fase le risorse si redistribuiscono tra apprendimento lessicale e sviluppo del sistema grammaticale.

Larsen-Freeman (1997; 2002; 2006; 2007), riprendendo il lavoro di van Geert (1991; 1994), utilizza l'approccio emergentista come utile prospettiva per spiegare due aspetti chiave dell'apprendimento della L2: la possibilità di ottenere diversi *pattern* di sviluppo a livello individuale e l'interpretazione del ruolo della variazione. Catturare il dinamismo dell'interlingua è un tema ampiamente discusso nelle ricerche acquisizionali. Diversi fattori infatti intervengono nel determinare il percorso di sviluppo, alcuni di questi esterni all'individuo, come la lingua da apprendere, la marcatezza della L1, la quantità e tipologia di input, la quantità e il tipo di interazione, la quantità di feedback, l'eventuale insegnamento, altri interni all'individuo, come età, fattori socio-psicologici quali motivazione e attitudine, caratteristiche individuali, stile cognitivo, ecc. Nessuno di questi elementi preso singolarmente giustifica il progresso, mentre la loro interazione ha effetti notevoli in quanto "both individual creativity and social interaction combine to influence the shape of the developing grammar" (Larsen-Freeman, 1997; cfr. meta-analisi di Goldschneider & DeKeyser, 2001). La lingua cresce e si auto-organizza in modo organico dando origine a una serie di strutture stabili, ma nel complesso poco è fisso e immutabile. Il termine stesso lingua target è ingannevole. Se la lingua è un sistema dinamico che si adatta continuamente al contesto, il target degli apprendenti è in costante movimento, non esiste una fine del percorso di acquisizione poiché non ci sono vincoli che definiscono a priori il punto di arrivo della traiettoria di sviluppo. L'apprendimento poi non procede linearmente, con gli apprendenti che interiorizzano completamente una struttura prima di passare alla successiva, ma piuttosto, così come tipico dei sistemi dinamici, "development is not discrete and stage-like but more like the waxing and waning of patterns; that, from a target-language perspective, certain aspects of the behaviour are progressive, others, regressive; that change can be gradual and it can be sudden" (Larsen-Freeman, 2006: 1; cfr. Thelen & Bates, 2003: 381)". Non ci sono stadi discreti in cui l'interlingua non è variabile, nonostante possano esserci periodi in cui alcune strutture sono predominanti: "the dominance of certain structures may arise through a gradual

building up process or through a period of fluctuation among competing forms, followed by a shift in the system when a certain critical threshold is crossed, and some wider reorganisation is triggered” (Larsen-Freeman, 2006: 592). La non linearità è data dal fatto che i diversi sotto-sistemi sono interconnessi: in alcuni casi supportano uno lo sviluppo dell’altro, in altri entrano in competizione. Questo accade perché le risorse nei sistemi dinamici sono per definizione limitate, di conseguenza una miglior performance in una dimensione di proficiency può sottrarre risorse alle altre dimensioni.

Secondo l’approccio emergentista, la variazione nella performance in quanto proprietà intrinseca dei sistemi dinamici di un apprendente è riscontrabile sia a livello sincronico che diacronico: non è un errore di misurazione o rumore da eliminare, ma piuttosto il risultato della capacità di sviluppo e di adattamento del sistema: essa è sia origine dello sviluppo che indicatore di un momento specifico del processo di sviluppo stesso. La psicolinguistica, la sociolinguistica, la linguistica dei corpora, i sistemi di analisi probabilistica (*probabilistic analyses*) giustificano tale instabilità come sistematica e governata da regole (per es. Labov, 1970; cfr. § 2.2.2.1), l’approccio emergentista interpreta la sistematicità della variazione in modo dinamico rafforzando il senso dell’espressione ‘competenza variabile’:

Competence and performance both emerge from the dynamic system that is the frequency-tuned conspiracy of memorized exemplars of construction usage, with competence being the integrated sum of prior usage and performance being its dynamic contextualized activation. (Ellis & Larsen-Freeman, 2006; cfr. Ellis, 2006a; 2006b; 2006c).

I percorsi individuali di sviluppo possono essere molto diversi l’uno dall’altro, anche se visti nell’insieme possono apparire piuttosto simili: gli apprendenti differiscono nelle loro abilità d’uso della lingua non perché hanno o non hanno una conoscenza specifica, ma piuttosto le differenze sono il risultato della variazione nei modi e nei tempi in cui il processo è stato realizzato durante l’apprendimento (cfr. Larsen-Freeman, 2006; de Bot *et al.*, 2007). Larsen-Freeman (2006) ad esempio esaminando i dati di cinque cinesi apprendenti di inglese L2 rileva proprio come vi sia un’evidente variazione a livello

individuale nei percorsi di sviluppo. L'analisi qualitativa del suo corpus evidenzia infatti come l'apprendimento proceda come in una spirale, con forme complesse che emergono, vengono abbandonate e poi riprese con maggior accuratezza in una fase successiva. Utilizzare un approccio emergentista richiede dunque di concentrarsi non tanto o non solo sulle tendenze generali a partire da dati aggregati a livello di gruppo, ma anche sull'analisi di diversi casi individuali osservati longitudinalmente, così da ricostruire cosa accade realmente durante il processo di apprendimento. Dati di gruppi e dati individuali risultano così complementari:

individual data provide us with a window into the complex world of interacting vincolis in the learner. Group data, on the other hand, provide us with a window toward society and the distribution of the linguistic input (MacWhinney, 2006: 738).

Dal punto di vista metodologico, un approccio di questo tipo pone diverse sfide alle ricerche future. Come suggeriscono Ellis & Larsen-Freeman (2007) uno studio appropriato dei costrutti linguaggio e apprendimento linguistico richiede la sinergia di diversi approcci metodologici:

Each research methodology has its advantages. Ethnography brings together the individual, society, and consciousness in time and place, but it ignores implicit motivations that introspection cannot access (Nisbett & Wilson, 1977). Brain imaging illustrates the dynamic patterns of neural activity involved in mental processing, but it isolates the learner's brain from society and its normal ecology of function. Laboratory experimentation allows the controlled logic of the scientific method, but it sacrifices ecological validity in this goal. And so on. We need to bring these methods together, to strive after the linking relations, to develop the kind of account that Wertsch, a Socioculturalist, called 'translation at the crossroads that would make it possible to link, but not reduce, one perspective to another' (Wertsch, 1998) (Ellis & Larsen-Freeman, 2007: 577-78)

Lo studio dell'acquisizione secondo l'approccio emergentista richiede l'impiego di specifici strumenti: sarà necessario predisporre corpora longitudinali rappresentativi di una molteplicità di condizioni, in modo da osservare i fenomeni di sviluppo in un contesto più ampio (cfr. van Geert & Steenbeek, 2005), e selezionare strumenti di misurazione capaci sia di catturare l'intero processo di sviluppo (cfr. van Geert & van Dijk, 2002; Larsen-Freeman, 2006) sia essere sensibili alle variazioni individuali (cfr. Ellis & Larsen-Freeman, 2007).

L'emergentismo rappresenta un nuovo approccio nello studio del linguaggio e certamente offre una serie di metafore utili a descrivere la complessità del processo di apprendimento linguistico, metodi per fare predizioni o interpretare complesse interazioni dialogiche. Nonostante le potenzialità di tale approccio, alcuni ricercatori ritengono che i sostenitori dell'emergentismo siano anche troppo ottimisti (cfr. Horgan, 1995; MacWhinney, 2006; Cameron & Deignan, 2006; Pienemann, 2007; Stephan, 1999) e mettono in luce alcune limitazioni da tenere in considerazione. Pienemann (2007: 43) ad esempio evidenzia come questo approccio da un lato non prenda in sufficiente considerazione le ricerche che lo hanno preceduto, dall'altro non definisca in modo chiaro cosa si intende per variazione: manca una definizione nuova e operativa di variazione o più in generale delle dinamiche dello sviluppo di una L2, capace di offrire una valida alternativa agli approcci consolidati usati in letteratura. MacWhinney (2006), pur sostenendo l'utilità dell'emergentismo, invita alla cautela, pone alcune questioni metodologiche e si chiede come l'emergentismo possa raggiungere un livello soddisfacente di applicazione empirica negli studi acquisizionali: è abbastanza semplice elaborare attraenti interpretazioni emergentiste, ma è altrettanto necessario verificare empiricamente che non siano errate. Una volta poi stabilito che l'apprendimento è un sistema dinamico in cui i diversi elementi sono emergenti sarà necessario definire cosa emerge e soprattutto come. Anche MacWhinney (2006) sottolinea proprio come lo sviluppo futuro di questo approccio è strettamente legato alla creazione di adatti supporti metodologici, legati in particolare alla linguistica dei corpora e al potenziamento dei metodi di analisi computazionale. In ragione di questi limiti, non adotterò un approccio emergentista in questo studio. Ho ritenuto comunque utile presentare questa parte di letteratura sia perché la più recente sul tema della variazione, sia per l'attenzione ai percorsi individuali, sia per l'utilità descrittiva di alcune delle metafore impiegate.

Ricapitolando, la letteratura acquisizionale ha affrontato il tema della variazione dell'interlingua a partire da diversi approcci: quelli di tipo sociolinguistico, che indagano principalmente l'ordine di emersione o la qualità e variazione d'uso di strutture specifiche in relazione a contesto linguistico e situazionale; quelli di tipo psicolinguistico, che tentano di descrivere la proficiency e lo sviluppo attraverso misure quantitative delle dimensioni della performance; e infine il più recente approccio emergentista, che inquadra la variazione entro le teorie dei sistemi complessi. Gli studi condotti hanno dunque contribuito in modo complementare a una maggior comprensione del ruolo e dell'importanza della variazione nell'apprendimento. Questo studio pertanto, utilizza l'approccio psicolinguistico come ancoraggio teorico e metodologico principale, mentre l'approccio sociolinguistico motiva l'attenzione alla variazione situazionale e al confronto con il comportamento dei parlanti nativi, e infine l'approccio emergentista giustifica l'attenzione per la variazione individuale in relazione a quella di gruppo.

#### 2.2.2. La misurazione della variazione

Nella sezione precedente (§ 2.2.1.2) abbiamo visto che ricerche come quelle di Skehan (1998) e Robinson (2001) dimostrano come l'apprendente possa avere obiettivi diversi nello svolgere un compito linguistico, favorendo pertanto di volta in volta una delle tre dimensioni CAF. L'utilità dell'analisi di queste tre dimensioni è motivata sulla base di modelli psicolinguistici (§ 2.1.2) e sostenuta dai risultati ottenuti in vari studi (§ 2.2.1.1). L'interesse per l'analisi CAF è condiviso anche dalle ricerche sugli indici di acquisizione (*developmental index studies*) interessate all'andamento longitudinale di queste tre dimensioni e delle loro relative misure o al legame con il livello di proficiency. Sin dalla fine degli anni '70 diversi studi hanno tentato di individuare una misura oggettiva dello sviluppo linguistico nella seconda lingua, "a developmental yardstick against which global (i.e. not skill or item



specific) second language proficiency could be gauged” (Larsen-Freeman, 1983: 287). Misure di questo tipo sarebbero infatti utili a insegnanti, tester e ricercatori. Dal punto di vista dell’insegnamento permetterebbero di accertare gli effetti di un intervento educativo o di una tecnica didattica; dal punto di vista del testing costituirebbero un utile strumento per definire il profilo di competenza, e per validare prove e scale olistiche. Dal punto di vista della ricerca, renderebbero più affidabile il confronto tra studi diversi e tra lingue diverse.

Negli ultimi anni la letteratura relativa alle misure quantitative CAF si è dimostrata piuttosto vivace. Ad esempio, con riferimento alla lingua scritta, Polio (1997; 2001) presenta criticamente un buon numero di misure impiegate in diversi studi, Wolfe-Quintero *et al.* (1998) verificano la *sampling reliability* e la *cuncurrent validity* di circa 100 misure CAF impiegate in diversi studi e Ortega (2003) nella sua meta-analisi approfondisce l’utilità di alcune di esse come misure di proficiency, con riferimento alla lingua orale. Freed (2000) offre un’interessante discussione sulle misure di fluenza, Ortega (1999), Skehan (2003) e Robinson & Ellis (2008) discutono invece della maggiore utilità di misure globali o misure specifiche.

Pubblicazioni di questo tipo offrono indicazioni utili sulle misure più affidabili, ma evidenziano anche alcuni limiti metodologici nelle ricerche condotte (e di conseguenza nelle rispettive analisi cumulative), in particolare in riferimento alle modalità di codifica dell’interlingua e all’uso di definizioni non univoche delle misure impiegate. Poiché le misure CAF dovrebbero aiutare il ricercatore a osservare con coerenza e affidabilità i propri dati, fornire una base teorica per l’interpretazione dei valori ottenuti e consentire la replica affidabile di uno studio, in letteratura è fortemente sentito il bisogno di migliorare comparabilità, affidabilità e validità delle misure stesse. In particolare due sembrano essere gli aspetti problematici su cui è necessario riflettere: da un lato complessità, accuratezza e fluenza sono costrutti complessi, di conseguenza è necessario che le misure impiegate siano in grado di dar conto delle diverse sotto-componenti delle tre dimensioni, dall’altro sono

per natura legate allo sviluppo, se quest'ultimo è un sistema dinamico fatto di diversi fenomeni che interagiscono tra di loro in modi non prevedibili, le misure impiegate dovranno riuscire a riflettere anche questo aspetto.

Nonostante le debolezze evidenziate, autori come Wolfe-Quintero *et al.* (1998) sottolineano come proprio le potenzialità di tali misure giustifichino l'interesse per ulteriori indagini in questa direzione e identificano alcuni obiettivi per la ricerca sulle misure CAF: la determinazione della migliore unità di produzione tra clausola o frase complessa, la determinazione della linearità o non-linearità nell'andamento delle misure, il rapporto tra le singole misure CAF e i relativi costrutti, il rapporto tra misure e livelli di competenza, e tra misure e sequenze di acquisizione. Il presente studio cerca di offrire un contributo in questa direzione. Le sezioni che seguono illustrano le unità utilizzate in letteratura per segmentare i testi parlati e per quantificare CAF nelle produzioni.

#### 2.2.2.1. L'unità di produzione

Un'unità di segmentazione del parlato è uno strumento essenziale se l'intento è analizzare aspetti quantitativi della competenza linguistica quali complessità, accuratezza e fluenza, o studiare elementi che caratterizzano il profilo di un apprendente, l'evoluzione nell'apprendimento o la variabilità situazionale dell'interlingua. Se tra gli obiettivi rientrano poi la comparabilità tra studi e tra lingue, o la meta-analisi, allora diventa fondamentale concordare sulla natura dell'unità, sulla sua applicabilità a diversi tipi di produzione nonché accertarsi che sia sufficientemente valida e affidabile.

La segmentazione di testi orali infatti non è scontata, e apre piuttosto alcune questioni teoriche e metodologiche interessanti rispetto ad almeno tre aspetti della sintassi nella conversazione: la definizione di clausola e frase; la classificazione delle relazioni tra clausole e l'interazione tra prosodia e sintassi.

Il parlato è suddiviso in proposizioni, ma l'identificazione delle

proposizioni, dal punto di vista sintattico o semantico, non è scontata (cfr. Sornicola, 1981; Miller & Weinert, 1998). In particolare le proposizioni parlate non coincidono sempre con la clausola, e di conseguenza vi è il rischio di una certa arbitrarietà nella suddivisione del continuum parlato in unità (cfr. Moneglia & Cresti, 1998; Moneglia, 2000a). La linguistica moderna, concentrando l'attenzione principalmente sulle strutture sintattiche complesse, ha parzialmente trascurato alcuni elementi sintattici che hanno scarsa struttura ma grande importanza nello scambio enunciativo tra i parlanti. Le definizioni tradizionali di clausola come stringa contenente una predicazione, o di frase come espressione di senso compiuto, non soddisfano le richieste di analisi del parlato. Consideriamo a titolo esemplificativo gli estratti (1) e (2) dal corpus oggetto di studio di questa tesi:

- (1) 26 ELI: || { .hh poi:: mh:: tch e:: #2\_4 hh e:: } spetta. | ||  
 27 ELI: || { fhhf } un attimo devo:  
 28 INV: eheheh tranquilla tranquilla  
 29 ELI: mettere a posto tutte le:: | ||  
 30 INV: le se[quenze  
 31 ELI: || { [eh. ] poi sì. | ||
- (2) 48 INV: volevo semplicemente sentire:: da te [un un racconto+  
 49 ELI: || [comunque mh:: #0\_2 carino | || #0\_6  
 50 INV: +degli eventi che succedono  
 51 ELI: || carino. | ||  
 52 ELI: || cioè # bello. | ||  
 53 ELI: || che poi si incontrano. | ||

In questi esempi vi sono almeno quattro diverse strutture sintattiche che mettono in crisi l'applicazione della definizione tradizionale di frase: (a) frammenti (1.29; 2.48); (b) clausole senza verbo, o con verbo sottinteso o presupposto (1.28; 1.30; 2.49; 2.51; 2.52); (c) unità sintattiche distribuite su più turni (1.27-1.29; 2.48-2.50); (d) clausole dai confini incerti (2.52-2.53).

Se questi elementi non vengono definiti in modo inequivocabile, possono generare arbitrarietà nell'interpretazione. Per questo motivo alcuni ricercatori suggeriscono addirittura di eliminarli dall'analisi (per es. Biber *et al.*, 1998). Un approccio di questo tipo avrebbe però il grosso limite di considerare come frammenti e quindi escludere dall'indagine scientifica circa il 40% delle locuzioni (Moneglia, 2000b; 2000c).

D'altra parte, poiché non pare possibile proporre una nuova definizione di frase che possa superare questi problemi (cfr. Graffi, 2001), la letteratura ha identificato due soluzioni: utilizzare un approccio pragmatico e individuare tutte le strutture che possono esprimere un senso compiuto, oppure utilizzare un approccio sintattico e distinguere tutte le combinazioni di parole che possono essere identificate come frasi. Nei paragrafi che seguono presenterò alcuni approcci del secondo tipo.

Per l'italiano L1 è particolarmente rilevante il lavoro condotto da Cresti e dai suoi collaboratori (cfr. Cresti, 1987; 1992a,b; 1993; 1994; 1995; 2001; 2004; 2005; Cresti & Gramigni, 2004; Moneglia & Cresti, 1997; Cresti & Moneglia, 2005; Moneglia, 1999; 2000a; 2000b; 2004). Cresti propone una chiara distinzione tra frase ed enunciato. In contrapposizione con l'approccio generativista, che identifica l'enunciato come entità dell'esecuzione e la frase come entità basica della competenza, qui enunciato e frase sono visti come entità di riferimento di varietà diamesiche diverse (parlato vs scritto). Ciascuna delle due varietà sarà costituita da frasi e da enunciati, e ciò che le distingue è il peso che danno a queste due diverse entità: nel parlato infatti il livello caratterizzante non è tanto quello logico-sintattico della frase, quanto quello pragmatico-illocutivo e dell'articolazione dell'informazione, fondato appunto sull'enunciato. La ricercatrice propone un metodo molto dettagliato per identificare l'enunciato, e cioè l'unità linguistica superiore alla parola con 'valore di comunicazione'. La novità di questo approccio risiede nel fatto che la segmentazione dei testi in enunciati si basa sull'analisi degli indizi prosodici con il loro valore funzionale. L'enunciato viene definito come

ogni espressione linguistica interpretabile pragmaticamente, legata: a) ad una condizione semantica di piena significanza dell'espressione in questione (parola lessicale vs morfema); b) della sua realizzazione intonata secondo un pattern melodico di valore illocutivo. Secondo tale proposta l'enunciato è l'unità di riferimento della lingua parlata; esso costituisce il corrispettivo linguistico di un atto ed è prosodicamente identificabile nel continuum parlato (criterio illocutivo e articolazione informativa). In tal modo l'enunciato risulta definito dal carattere dell'autonomia, ovvero dell'interpretabilità pragmatica, e l'indice linguistico necessario alla sua realizzazione è l'intonazione. Asseriamo che l'autonomia, o interpretabilità pragmatica, è proprietà diversa dalla

compiutezza semantica, nè pare vincolata alla realizzazione di una predicazione, in particolare a quella costituita da una forma verbale finita, come invece previsto per la proposizione.” (Cresti, 2005: 2)

Altri autori hanno cercato di risolvere il problema dell’identificazione delle unità del parlato e la letteratura a proposito è piuttosto vivace. La tabella 2 riporta le principali strutture utilizzate nelle ricerche sulle L2 (per una presentazione dettagliata, per l’inglese: Crookes, 1990, Foster *et al.*, 2000; Ellis & Barkhuizen, 2005; per l’italiano L1: De Mauro & Thorton, 1985; Voghera, 1992; Cresti, 2005; per l’italiano L2: Chini, 1998; 2003; Valentini, 1988; 2001)

**Tabella 2. L’unità di produzione**

unità sintattiche	
<i>T-unit</i>	<i>minimal terminal unit</i> , formata dalla principale e dalle subordinate a essa collegate; è ampiamente impiegata nelle ricerche sulle produzioni scritte (Hunt, 1965; 1966; 1970)
<i>S-node</i>	“indicated by tensed or untensed verbs” (Ellis <i>et al.</i> , 1994: 483)
proposizione	unità semantica che consiste di almeno un argomento e uno o più predicati (Sato, 1988: 375)
enunciato	una stringa di parlato con almeno una delle seguenti caratteristiche: (a) sotto uno stesso profilo intonativo; (b) circondata da pause; (c) contenente un’unica unità semantica, e.g. Crookes (1990: 187)
<i>C-unit</i>	<i>communication unit</i> , consiste di predicazioni grammaticali indipendenti o risposte cui manca la ripetizione di un elemento della domanda per soddisfare il criterio di indipendenza, Loban (1966, in Ellis e Barkhuizen, 2005), e di proposizioni, parole, sintagmi o clausole, grammaticali e agrammaticali, che hanno un significato referenziale o pragmatico, secondo Pica <i>et al.</i> (1989: 72)
<i>AS-unit</i>	proposizione prodotta da un solo parlante fatta di una clausola indipendente o di una unità indipendente inferiore alla clausola e tutte le clausole subordinate a essa associate (Foster <i>et al.</i> , 2000)
altre unità	
turno	misura applicabile a situazioni dialogiche: “stream of speech bounded by speech of another, usually an interlocutor” (cfr. Sacks <i>et al.</i> , 1974; Ford <i>et al.</i> , 1996; Selting, 1996)
<i>unità tonale</i>	definita da una configurazione distinta di pitches con una chiaro centro e nucleo. Il nucleo è la sillaba (o in alcuni casi una serie di sillabe) che ha la maggior preminenza nell’unità tonale (Crystal e Davy, 1975: 16)
<i>unità di idea</i>	un segmento di messaggio costituito da topic e comment, separato sintatticamente o intonazionalmente dalle unità contigue. Serve per il conteggio del numero di idee espresse in un testo (Kroll, 1977) e misura la completezza proposizionale, fornendo una misura del grado di concettualizzazione raggiunto dall’apprendente (cfr. Levelt, 1989). Viene impiegata per produzioni di contenuto predefinito, quali quelle elicitate da un task di racconto di una storia (cfr. Chafe, 1980; Danielewicz, 1984; Larsen-Freeman, 2006)

In ogni caso l'unità più adatta per la segmentazione del parlato deve essere valida e affidabile, due qualità alla base di qualsiasi strumento di misurazione. Una misura è affidabile se dà la stessa lettura dello stesso testo orale in momenti diversi; ed è valida se misura ciò che intende misurare. Nel testing l'affidabilità di un sistema di misurazione è solitamente valutata in termini di affidabilità del giudizio tra valutatori (*inter-rater reliability*). La validità dello strumento ha invece a che vedere con il legame che intrattiene con i processi psicologici legati alla produzione linguistica. Come nota Crookes (1990), se la misura non corrisponde a processi di pianificazione psicolinguistica, allora la misurazione effettuata avrà scarso o nessun valore.

Nel caso delle misure sopra descritte manca spesso un'attenzione per la loro validazione. Almeno tre elementi limitano infatti l'applicazione a nuove ricerche delle unità impiegate: (a) la definizione: unità uguali vengono chiamate in modo diverso nei vari studi, spesso con definizioni imprecise, ma anche quando le descrizioni sono precise, di solito sono troppo semplici per rispondere alle esigenze dei dati reali; (b) gli esempi: se esemplificate, le definizioni sono accompagnate da uno o due esempi soltanto, e in genere poco rappresentativi del parlato, in particolare di quello interattivo; (c) l'affidabilità: raramente le ricerche riportano un'analisi dell'affidabilità dell'applicazione della misura.

Questi punti di debolezza riducono l'utilità di molte di queste misure. Foster *et al.* (2000) cercano di superare i limiti delle unità esistenti proponendo l'*Analysis of Speech Unit* (d'ora in poi AS-unit). Questa unità è principalmente sintattica, piuttosto che intonativa o semantica, anche se intonazione e pause sono comunque prese in considerazione nell'interpretazione di casi dubbi. Foster *et al.* (2000) definiscono la AS-unit come "a single speaker's utterance consisting of an *independent clause*, or *sub-clausal unit*, together with any *subordinate clause(s)* associated with either. [...] False starts, functionless repetitions, and self-correction are put inside brackets". Distinguono poi le clausole che costituiscono le AS-unit in clausole indipendenti, unità

indipendenti inferiori alla clausola, e clausole dipendenti.

Come suggeriscono gli autori, l'AS-unit sembra avere una serie di vantaggi. Primo, dal punto di vista psicolinguistico, la sua validità è sostenuta dalle ricerche condotte su pause e pianificazione: le unità sintattiche sono anche unità di pianificazione (per es. Raupach, 1980; Garman, 1990), i parlanti poi pianificano unità fatte di più clausole (Beattie, 1980). Secondo, le unità sintattiche sono più affidabili in quanto più facilmente identificabili rispetto a unità intonative o semantiche (Foster *et al.*, 2000). Infine, come sostengono Foster *et al.*, l'AS-unit è applicabile a una gran varietà di dati orali. Sulla base di queste considerazioni ho scelto la AS-unit come unità di produzione per la mia analisi. L'applicazione di tale unità ai dati del mio corpus ha richiesto una ridefinizione di alcuni aspetti. Per la discussione sulla sua applicazione all'italiano si veda § 3.4.

#### 2.2.2.2. Le misure analitiche

Abbiamo visto in precedenza come complessità, accuratezza e fluenza rappresentino diverse dimensioni della performance. La complessità è vista in genere come capacità di produrre strutture varie e complesse, l'accuratezza è data dall'assenza di errori nelle produzioni e infine la fluenza è legata alla velocità. Quando osserviamo complessità e accuratezza osserviamo gli effetti di *representation* e *restructuring*, quando osserviamo la fluenza vediamo l'automaticità delle produzioni (§ 2.1.2).

La misurazione CAF dell'uso linguistico dovrebbe dunque permettere riflessioni sui processi di apprendimento, nonché aiutare a distinguere studenti a diversi livelli di proficiency. Larsen-Freeman (1978: 440) in un articolo ormai di diversi anni fa sottolinea come una misura ideale per l'abilità linguistica dovrebbe "increase uniformly and linearly as learners proceed towards full acquisition of a target language". Tale preferenza sembra giustificata dalla pratica didattica e dal testing: in genere le scale di proficiency

sono lineari. Come riconoscerà la stessa Larsen-Freeman (1997), l'attesa che il ritmo di sviluppo sia uniforme in uno stesso individuo e nel gruppo non è stata giustificata dalle ricerche condotte da allora ad oggi. Ad esempio il ritmo varia sostanzialmente da un individuo all'altro (per es. Perdue & Klein, 1993; Skehan, 1989; Slavoff & Johnson, 1995; Larsen-Freeman, 2006) e il legame tra lo sviluppo di una proprietà x e il conseguente sviluppo di una proprietà y, tipico dell'apprendimento della L1, non è sempre riscontrabile nell'acquisizione della L2 (per es. Clahsen & Muysken, 1986; 1989; Meisel, 1997). Ne consegue che la richiesta di un indice che cresca linearmente non è né necessaria né tantomeno giustificata. Un obiettivo più realistico sarebbe identificare un gruppo di misure che, combinate insieme, produca un profilo dell'apprendente che può poi essere classificato in modo affidabile a un dato livello di una scala predeterminata.

Nel selezionare tale gruppo di misure, così come abbiamo visto per l'unità di segmentazione del parlato, dobbiamo sottometterci ai criteri della validità e affidabilità. Wolfe-Quintero *et al.* (1998: 117-18) e Norris & Ortega (2003) hanno individuato tre fonti di informazioni che permettono di valutare la *construct validity* di una particolare misura di L2 proficiency: validità predittiva (*predictive validity*), *concurrent validity*, e *repeated sampling reliability*. Validità predittiva si otterrebbe se gli stadi di sviluppo di un particolare aspetto della grammatica avessero valori corrispondenti a quelli della misura in questione. Questa era stata ad esempio la motivazione iniziale che aveva spinto a considerare la lunghezza media dell'enunciato come misura dello sviluppo linguistico nei bambini (Brown, 1973). La *concurrent validity* si ottiene invece nei casi in cui la misura correla con differenze tra livelli di proficiency o con i livelli di proficiency o con entrambi. Infine, la *repeated sampling reliability* si otterrebbe quando una misura dimostra un andamento lineare tra i diversi livelli di proficiency. Wolfe-Quintero *et al.* (1998) notano come la validità predittiva delle misure non è stata indagata dai ricercatori di SLA e attraverso la loro meta-analisi si concentrano sulla valutazione di *concurrent validity* e *repeated sampling reliability*. L'affidabilità delle misure



potrà invece essere ottenuta tramite la formazione dei valutatori e il double-marking. Quest'ultimo tipo di verifica sull'affidabilità è raramente applicato in letteratura.

Wolfe-Quintero *et al.* (1998: 119) attraverso la loro meta-analisi di misure CAF indicano quali tra quelle utilizzate nella ricerca valga la pena implementare. Nelle sezioni che segue discuto le diverse misure CAF proposte dalla letteratura, in § 3.4.1 presento le misure selezionate per questo studio.

La prima dimensione è la complessità. Foster & Skehan (1996) definiscono lo sviluppo della complessità come “progressively more elaborated language” e “a greater variety of syntactic patterning” (p. 303). Tale costrutto è rilevante in SLA poiché implica che l'apprendimento linguistico richieda, tra altri processi, la crescita del repertorio sintattico in L2 e l'abilità nell'utilizzarlo in modo appropriato nelle diverse situazioni comunicative (§ 2.1.2). Inoltre le validazioni a posteriori dei test linguistici provano come la grammatica sia un fattore significativo nel differenziare tra punteggi o tra livelli di proficiency globale (cfr. Ellis, 2001). La complessità sintattica ha a che fare con la gamma di strutture sintattiche impiegate e riguarda la possibilità per l'apprendente di accedere rapidamente a strutture sia semplici che sofisticate. L'analisi della complessità sintattica dovrebbe quindi rilevare quanto varie o complesse sono le unità di produzione, o il lessico o determinate strutture grammaticali e discorsive: lunghezza dell'unità di produzione, quantità di subordinazione e gamma di tipi di strutture sono gli elementi presi in considerazione dalle diverse misure analitiche presenti in letteratura.

Le misure oggettive della complessità sintattica devono il loro sviluppo in buona parte agli studi sulla L1, in particolare al lavoro di Hunt (1965; 1970), che occupandosi di produzioni scritte ha dimostrato come la lunghezza della frase (individuata sulla base della punteggiatura) non sia un indicatore sufficiente della proficiency linguistica, poiché non vi è sempre corrispondenza tra livello di proficiency e lunghezza della frase. Hunt propone quindi un'unità di produzione alternativa, la *minimal terminal unit*, T-unit (cfr. § 2.2.2.1), definita come clausola principale e tutte le clausole a essa collegate. La

lunghezza media delle T-unit (MLU), combinata con misure che calcolano l'incidenza di clausole subordinate e indipendenti, offre indicazioni affidabili sulla maturità sintattica delle produzioni scritte di apprendenti di L1.

Misure analoghe sono state proposte anche per la L2, dove la complessità sintattica viene misurata attraverso il calcolo della lunghezza media dell'unità di produzione (per es. AS-unit) o con misure di tipo grammaticale che rilevano il grado di subordinazione. Quest'ultime sono di due tipi: misure che analizzano la clausola, la T-unit o la AS-unit in rapporto l'una all'altra (per es. parole/AS-unit; clausole dipendenti/AS-unit) e misure che analizzano la distribuzione di una determinata struttura sintattica (per es. passivi/AS-unit) (cfr. Foster & Skehan, 1996). Wolfe-Quintero *et al.* (1998) attraverso la loro meta-analisi concludono come lunghezza media di T-unit o clausola, clausole per T-unit e clausole subordinate per T-unit o clausole correlino bene con gruppi di livello e valutazioni olistiche. La tabella 3 riassume le principali misure impiegate.

**Tabella 3. Le misure CAF della complessità sintattica**

<i>Lunghezza</i>	lunghezza media della clausola, lunghezza media della T-unit/AS-unit. Misura proposta per la L1, frequentemente usata in SLA (per es. Brown, 1973; Hunt, 1965; Loban, 1976; Scott, 1988)
<i>Grado di subordinazione</i>	clausole dipendenti per T-unit/AS-unit, clausole dipendenti per clausole, clausole dipendenti per clausole complete (per es. Elder & Iwashita, 2005; Skehan & Foster, 2005; Michel <i>et al.</i> , 2007)
<i>Grado di coordinazione</i>	misura di coordinazione, misura suggerita da Bardovi-Harlig (1992) per i livelli elementari
<i>Frequenza d'uso di alcune strutture</i>	conteggio della frequenza di alcune strutture della morfologia verbale (per es. passivi), della classi di verbi (per es. modali) o di strutture sintattiche (per es. relative), cfr. Ellis & Yuan (2005), Robinson (2007b)

Quando si selezionano le misure da impiegare in uno studio è necessario stabilire di quante e quali di queste misure si ha bisogno per rilevare correttamente il costrutto complessità: non tutte le misure sopra presentate sono tra loro equivalenti, alcune misure potrebbero essere ridondanti se utilizzate contemporaneamente, altre potrebbero essere complementari, in quanto rilevano aspetti diversi del costrutto.

Vediamo più in dettaglio quali diversi aspetti della complessità sintattica dovrebbero essere colti dalle misure impiegate. È generalmente condiviso in letteratura che la complessità sintattica si sviluppi in due modi: un apprendente può realizzare clausole più complesse, aggiungendo sintagmi e parole, oppure può produrre unità più complesse, realizzando subordinate. Poiché misurare la complessità sintattica significa osservare questi due aspetti è necessario chiedersi che tipo di complessificazione viene rilevata dalle diverse misure. Considererò in particolare lunghezza media dell'unità di produzione e grado di subordinazione mettendo in evidenza la loro complementarietà. Non considererò qui in dettaglio le misure specifiche legate a un numero limitato di strutture ritenute complesse impiegate in alcuni degli studi più recenti, poiché questo tipo di misura non verrà impiegata in questo studio.

Le misure di lunghezza sono due: lunghezza media dell'unità di produzione e lunghezza media della clausola. La prima è ampiamente usata negli studi sullo sviluppo delle competenze di produzione scritta. Un aumento della lunghezza media della T-unit è generalmente considerato un incremento della complessità sintattica. La misura lunghezza media della T-unit (numero di parole per T-unit), come sottolinea lo stesso Hunt (1965), coincide con il numero di parole per clausola moltiplicato per il numero di clausole per T-unit, di conseguenza la misura parole per T-unit (o AS-unit) può aumentare sia come conseguenza della produzione di clausole più lunghe che di una maggior quantità di subordinate. Sulla base di queste osservazioni risulta chiaro come la misura della lunghezza media della AS-unit potrebbe essere una misura globale della complessità sintattica, ma non permette certo di distinguere la qualità della complessificazione. Consideriamo ora invece la misura lunghezza media della clausola. Il suo incremento nel numero di parole è il risultato dell'aggiunta di modificatori o dell'uso di strutture complesse, come nel caso della nominalizzazione. Questa misura rileva dunque l'aumento della complessità a livello di clausola. Le due misure di lunghezza sono spesso usate in letteratura come equivalenti, ma la discussione appena riportata mette in luce come le due misure siano in realtà diverse (cfr. la discussione in Norris &

Ortega, in stampa). Il conteggio delle parole per AS-unit può essere considerato una misura globale, il conteggio delle parole per clausola rileva invece un aspetto specifico della complessificazione. La lunghezza non è unanimemente considerata un efficace indicatore della complessità sintattica. Ad esempio, Wolfe-Quintero *et al.* (1998: 25), ammettendo di esprimere una posizione minoritaria, preferiscono addirittura considerare la lunghezza una misura di fluenza piuttosto che di complessità. Secondo gli autori questo è giustificato dalla constatazione che tale misura rileva direttamente solo la lunghezza, considerata da loro elemento della fluenza e non della complessità. A questo proposito è bene sottolineare come le due diverse interpretazioni possano essere entrambe giustificate dalla definizione stessa di AS-unit e di clausola. Queste infatti sono legate, oltre che alla sintassi, a fenomeni intonativi e pause. In questo senso possono assumere un doppio valore, riflettendo sia aspetti della fluenza che della complessità.

Vediamo ora le misure che rilevano il grado di subordinazione. Queste misurano la complessità sintattica in termini di aumento dell'uso della subordinazione. In letteratura sono stati proposti due tipi di misure: il calcolo delle clausole subordinate per unità di produzione e il conteggio delle clausole subordinate sul totale delle clausole. Entrambi le misure rilevano il grado di subordinazione e sono usati alternativamente o insieme nei diversi studi. Sulla base di queste considerazioni, in § 3.4.1 si verificherà sperimentalmente l'ipotizzata equivalenza tra le due misure. Le misure di subordinazione si sono dimostrate in generale efficaci nel distinguere livelli di proficiency (cfr. Wolfe-Quintero *et al.*, 1998; Ellis & Barkhuizen, 2005). Wolfe-Quintero *et al.* (1998) e Ortega (2003) nelle loro meta-analisi mettono comunque in luce alcune debolezze di queste misure e dei risultati che offrono. Da un punto di vista metodologico abbiamo già visto come il confronto tra i diversi studi sia limitato dal fatto che i diversi autori non sempre definiscono con precisione l'unità di produzione impiegata e non ne controllano validità e affidabilità. Dal punto di vista dell'andamento delle misure a diversi livelli di proficiency è

stato osservato come complessità sintattica in termini di aumento del grado di subordinazione e proficiency potrebbero non sempre correlare in modo lineare.

L'andamento non lineare del processo di complessificazione sintattica è stato osservato anche nel caso della L1. Scott (1988) ad esempio analizzando le produzioni scritte di alcuni alunni osserva come il grado di subordinazione si stabilizzi entro l'ottavo anno scolastico, mentre la lunghezza dell'unità di produzione continua a crescere, come risultato della complessificazione a livello di sintagma. Negli studi sull'acquisizione della L2 tale andamento non lineare è stato giustificato a partire da due ipotesi: l'Ipotesi di Predizione dello Sviluppo (*Developmental Prediction*) e l'Ipotesi di Transfer Retorico dalla L1 (*Cross-rhetorical transfer*).

L'Ipotesi di Predizione dello Sviluppo, proposta da Cooper (1976), Monroe (1975), Sharma (1980) e più recentemente da Wolfe-Quintero *et al.* (1998), sostiene una complessificazione non lineare della subordinazione, sulla base dell'assunto appunto che nella lingua scritta gli apprendenti avanzati tendano a complessificare a livello sintagmatico piuttosto che frasale. Da un punto di vista sperimentale tale predizione sarà supportata se la ricerca dimostrerà una relazione inversa tra grado di subordinazione e lunghezza delle unità di produzione ai livelli più alti. L'Ipotesi di Transfer Retorico dalla L1 è suggerita invece sulla base di uno studio condotto da Neff *et al.* (1998) in cui spagnoli apprendenti intermedio-avanzati d'inglese tendono ad avere valori di subordinazione sorprendentemente alti rispetto ai parlanti nativi dell'inglese. Ciò potrebbe dipendere da una maggiore tendenza a subordinare dello spagnolo scritto rispetto all'inglese, che verrebbe così trasferita alla seconda lingua. È probabile che entrambi i fattori, Ipotesi di Predizione dello Sviluppo e Ipotesi di Transfer Retorico dalla L1 possano giocare un ruolo importante nel determinare l'andamento dello sviluppo della complessità sintattica. Se ciò fosse vero, ci si dovrebbe attendere una graduale diminuzione da parte degli apprendenti a livello avanzato delle misure di subordinazione, una volta apprese le strategie retoriche appropriate della seconda lingua. Ne consegue che sarebbe probabilmente inappropriato considerare le misure di complessità

sintattica come misure assolute di sviluppo o indicatori diretti della competenza linguistica: un aumento o una diminuzione dei valori non corrisponde automaticamente a un cambiamento in positivo o in negativo, tutt'altro, una riduzione dei valori di subordinazione potrebbe essere comunque il risultato di un aumento della complessità globale della produzione. 'Più complesso' può significare più sviluppato in modi diversi e soprattutto non significa forzatamente meglio. Il progresso nella L2 può includere i due tipi di complessificazione sintattica, ma riguardare anche lo sviluppo di un repertorio discorsivo e sociolinguistico che il parlante può utilizzare in diverse situazioni comunicative. Se queste ipotesi sono corrette, il grado di subordinazione potrebbe essere una misura significativa ai livelli intermedi, mentre a livelli avanzati la lunghezza media della clausola potrebbe essere più rilevante. Si tratta di ipotesi suggestive, ma non ci risulta che nessuno studio abbia finora empiricamente dimostrato in modo sistematico questo tipo di flessione dei valori di subordinazione nei parlanti non nativi di livello più avanzato né indagato longitudinalmente lo sviluppo della complessità sintattica nel parlato di apprendenti avanzati. Attraverso l'analisi condotta in questo studio si cercherà di contribuire alla verifica sperimentale di ipotesi di questo tipo.

Riassumendo, la discussione appena illustrata dimostra come la complessità sintattica sia un costrutto multidimensionale, per misurarla efficacemente sarà utile quindi combinare diverse misure. Abbiamo visto come lunghezza media dell'unità di produzione può rappresentare una misura globale, mentre lunghezza media della clausola e grado di subordinazione offrono rispettivamente informazioni sulla complessificazione ottenuta attraverso l'elaborazione sintagmatica e quella dovuta all'aumento dell'uso della subordinazione, pertanto sono complementari. Le diverse misure del grado di subordinazione (cioè clausole subordinate/AS-unit e clausole subordinate/clausole) invece rilevano tutti la complessità in termini di aumento della subordinazione e si possono pertanto considerare tra loro equivalenti. Per l'applicazione di queste misure ai miei dati si veda § 3.4.

Passiamo ora alla seconda dimensione CAF, l'accuratezza. Foster & Skehan (1996) definiscono l'accuratezza come "freedom from error", ossia come uso target delle strutture linguistiche in contesto. In letteratura sono state impiegate diverse misure di accuratezza. Un primo tipo è rappresentato dal conteggio dell'incidenza di unità senza errori sul totale delle unità, un secondo tipo dal conteggio del numero e del tipo di errori.

La percentuale di unità senza errori sul totale delle unità come misura è stata criticata da alcuni studiosi. Il conteggio delle unità senza errori, dando lo stesso peso a un'unità di produzione contenente uno o più errori, non rivela molto sulla loro distribuzione, non distinguendoli in categorie non permette osservazioni sulla loro tipologia (Bardovi-Harlig & Bofman, 1989). Wolfe-Quintero *et al.* (1998) nella loro meta-analisi della validità e affidabilità delle diverse misure impiegate in letteratura mostrano però come il rapporto *unità di produzione senza errori sul totale delle unità* prodotte risulti essere una misura che meglio di altri correla con i progressi di breve periodo e con i giudizi olistici a diversi livelli di competenza. Queste osservazioni suggeriscono come una misura basata sul conteggio delle unità corrette possa essere un indicatore globale utile, anche se non esaustivo.

Alcuni ricercatori hanno proposto un secondo tipo di misura, basato sul conteggio del *numero* e del *tipo* di errori per proposizione o per clausola. In questo caso due sono le modalità di classificazione proposte, una basata sul tipo di errore – lessicale, morfo-sintattico, pragmatico o di pronuncia – l'altra basata sulla gravità dell'errore, definita in termini di comprensibilità. La classificazione degli errori per tipo ha messo in evidenza poi come, con il progredire della competenza linguistica, l'accuratezza non si sviluppi linearmente e come errori di diverso tipo possano avere evoluzioni più o meno lineari. Anche questa misura dimostra dunque alcuni punti di debolezza: nelle produzioni degli apprendenti non è sempre univoca l'assegnazione di un errore a una categoria, il conteggio indifferenziato di tutti gli errori inoltre non è sensibile alla quantità di contesti offerti. Nell'applicazione a dati spontanei e semi-spontanei si aggiungono difficoltà rispetto al controllo di una sufficiente

occorrenza di contesti obbligatori: alcune strutture potrebbero essere opzionali in quel determinato contesto comunicativo.

Un terzo tipo di unità cerca di superare questi ultimi limiti conteggiando e identificando le occorrenze corrette di determinate strutture. Le ricerche condotte in questa direzione si sono concentrate su un numero limitato di strutture. L'interesse qui è il possibile punto di contatto tra misure e studi sulle sequenze acquisizionali.

Wolfe-Quintero *et al.* (1998) nella loro meta-analisi evidenziano come per le misure di accuratezza considerate sopra non vi sia una chiara relazione con il livello di proficiency globale, ma piuttosto con valutazioni olistiche o progressi nel breve periodo. I risultati dei diversi studi sono spesso contraddittori. I ricercatori ipotizzano che le misure di accuratezza non siano collegate al costrutto sviluppo, ma piuttosto a quello di errore. Larsen-Freeman (1983) mostra come la maggior parte degli errori grammaticali riguardi aspetti della morfologia, articoli, preposizioni e tempi verbali, sostenendo che le strutture morfologiche risultino più complesse di quelle sintattiche. Bardovi-Harlig & Bofman (1989) dimostrano come sia gli studenti che hanno superato con successo il test che quelli che non l'hanno superato mostrano la stessa distribuzione di errori per tipo (57% errori morfologici, 27% errori lessicali, 15% errori sintattici), nonostante il gruppo meno competente commettesse in generale più errori. Lo studio suggerisce che possa esserci una differenza tra le strutture sintattiche più universali e centrali di una lingua, acquisite prima, e quegli aspetti più *language-specific*, periferici, locali della morfologia. Bardovi-Harlig & Bofman (1989: 28) definiscono questo come “strong syntax-weak morphology stage”. La definizione di come l'errore si relazioni allo sviluppo linguistico e la spiegazione delle ragioni per cui le misure di accuratezza (al contrario di complessità e fluenza) non correlano con i livelli di proficiency sono aspetti che la ricerca in futuro dovrà approfondire. Attraverso l'analisi condotta in questo studio si cercherà di offrire un contributo su questi temi. Nella sezione § 3.4 illustro le misure scelte in questo studio.

La terza dimensione CAF presa in considerazione è la fluenza. In ambito



acquisizionale il termine fluenza è interpretato in due modi. Secondo una definizione ampia il termine indica in generale la proficiency orale. In questo senso, con fluenza ci si riferisce al punto più alto di una scala che descrive la conoscenza della L2. Secondo una definizione più ristretta con fluenza ci si riferisce a una componente specifica della proficiency orale. Lennon (1990), riprendendo il lavoro di Mohle (1984), considera la fluenza in senso stretto, misurandola in termini di velocità e lunghezza della produzione. In questo senso la fluenza riguarda “the processing of language in real time” (Schmidt, 1992: 358) con attenzione primaria verso il significato (Foster & Skehan, 1996: 304). Essa può essere caratterizzata in due modi. Il primo modo riguarda l’uso appropriato di routine, dove per routine si intende sia una formula pragmatica (House, 1966) che espressioni automatizzate (Ellis, 1996). L’uso di routine genera un aumento nelle misure di fluenza in quanto facilita la realizzazione di unità di produzione. Il secondo modo riguarda l’automatizzazione dell’uso della lingua. Towell, Hawkins & Bazergui (1996: 85) affermano che “fluent second language production requires that knowledge (from whatever source) be proceduralised”.

Seguendo la definizione ristretta le misure di fluenza sono di due tipi (Wiese, 1984; Lennon, 1990): variabili temporali, in relazione alla velocità dell’eloquio e alle pause, fenomeni di esitazione in relazione alle disfluenze. Queste misure indicano il grado di successo di un apprendente nel gestire il *processing* in tempo reale tipico dell’interazione. Le variabili temporali si riferiscono alla velocità e lunghezza nella realizzazione di unità. I fenomeni di esitazione invece sono “connected to moment-by-moment decisions during performance, reflecting adjustments and improvements that are feasible within the pressure of real-time communication” (Skehan & Foster, 1999: 230). I fenomeni di esitazione possono essere collegati a differenze individuali, con apprendenti che monitorano costantemente le loro produzioni e studenti che non monitorano per nulla. Skehan (1998) dimostra attraverso un’analisi fattoriale l’indipendenza dei due tipi di misure di fluenza, definendo fluenza legata alla velocità (*speed fluency*) e fluenza legata alle pause (*breakdown*

*fluency*) le dimensioni della fluenza caratterizzata dalle variabili temporali e fluenza legata alle riparazioni (*repair fluency*) quella definita dai fenomeni di esitazione. Così come per la complessità sintattica (vedi sopra) il costrutto fluenza è multidimensionale. Un'analisi per essere efficace dovrà utilizzare misure rappresentative sia delle variabili temporali che di fenomeni di esitazione, così da riuscire a indagare le sotto-dimensioni della fluenza identificate. La tabella 4 riporta le principali misure impiegate distinguendole in rapporto alle dimensioni di fluenza: (a) fluenza legata alla velocità e (b) fluenza legata alle pause, entrambe rilevate con misure temporali; e (c) fluenza legata alle riparazioni basata sulle misure di auto-correzione e riformulazione.

**Tabella 4. Le misure di fluenza, adattato da Ellis & Barkhuizen (2005)**

<i>variabili temporali</i>	
<i>fluenza legata alla velocità</i>	
velocità d'eloquio	solitamente misurata in termini di numero di sillabe o parole prodotte al secondo o al minuto. Il numero delle sillabe disfluenze escluse ( <i>pruned syllables</i> ) e/o il numero delle sillabe disfluenze incluse ( <i>unpruned syllables</i> ) viene conteggiato e diviso per il numero totale di sillabe o parole prodotte
lunghezza tra due pause	numero medio delle sillabe o delle parole comprese tra due pause di una lunghezza predefinita (per es. 1 secondo). Questa misura non prende in considerazione le disfluenze
<i>fluenza legata alle pause</i>	
numero di pause	il numero totale delle pause, piene e vuote
durata delle pause	misurata in due modi: (a) la durata totale delle pause di una lunghezza predefinita (per es. 1 secondo); (b) la durata media delle pause di una lunghezza predefinita (per es. 1 secondo). La durata delle pause misura il silenzio durante un task
<i>fenomeni di esitazione</i>	
<i>fluenza legata alle riparazioni</i>	
false partenze	proposizioni incomplete (per es. frammenti) o piani abbandonati. Possono o meno essere seguite da riformulazioni
ripetizioni	parole, sintagmi o clausole ripetuti
riformulazioni	parole, sintagmi o clausole ripetute parzialmente o con auto-correzioni
sostituzioni	elementi lessicali sostituiti da altri elementi lessicali

Le misure della fluenza legata alla velocità più frequentemente usate in letteratura sono velocità d'eloquio e lunghezza tra due pause. Velocità d'eloquio è una misura che combina due aspetti temporali della produzione orale, il tempo di pianificazione *on-line* e la velocità di articolazione. La misura lunghezza tra due pause ci dà informazioni invece su quanto il parlante è in grado di produrre segmenti di messaggi senza pause. Questa misura rileva la capacità del parlante di accedere velocemente a 'pezzi' di lingua già pronti (sistema basato sugli esempi) e contemporaneamente osservare la capacità di organizzare il parlato per unità discrete (*unitization of spoken production*). Le misure solitamente impiegate per osservare la fluenza legata alle pause sono due, numero di pause e durata delle pause, e rilevano quanto il parlante deve interrompere il discorso per pianificare il messaggio. Gli apprendenti che fanno meno pause sono in questo senso più fluenti. Fluenza legata alla velocità e fluenza legata alle pause sono considerate indipendenti: un apprendente può ottenere una certa velocità d'eloquio e allo stesso tempo fare lunghe pause. Per i fenomeni di esitazione le misure sono quattro: false partenze, ripetizioni, riformulazioni e sostituzioni e indicano le capacità del parlante di sostenere efficacemente la comunicazione in tempo reale.

Wolfe-Quintero *et al.* (1998), ammettendo di esprimere una posizione minoritaria, preferiscono considerare lunghezza della T-unit o della clausola e lunghezza delle T-unit senza errori come migliori misure per la fluenza nella composizione scritta. La lunghezza media delle T-unit è interpretata come misura di fluenza piuttosto che di complessità sulla base della constatazione che la misura parole per T-unit non distingue le cause dell'incremento della lunghezza dell'unità di produzione, rilevando pertanto la lunghezza e non la complessità (§ sopra). La lunghezza della T-unit senza errori è in genere trattata come misura di accuratezza, proprio perché la correttezza è parte della formula. Wolfe-Quintero *et al.* (1998) suggeriscono come tale misura rileva la fluenza delle unità senza errori, proprio perché misura la lunghezza di tali unità, ma non il numero di errori o il grado di correttezza delle unità stesse (cfr. Ishikawa, 1995). La loro meta-analisi mostra come queste misure crescano

linearmente in relazione ai livelli di proficiency nei vari studi, indipendentemente da task, lingua target o definizione del livello di proficiency. L'unica eccezione è riportata nello studio di Larsen-Freeman (1983).

La letteratura sulla fluenza ha spesso messo enfasi sulla rapidità dell'eloquio dei parlanti nativi. È chiaro che la stessa fluenza dei nativi non è un costrutto uniforme e monolitico, ma vi è un certo grado di variabilità individuale legata all'argomento della conversazione, alla situazione, alle caratteristiche individuali del parlante o altri possibili fattori. A questo proposito Fillmore (1979: 93) distingue almeno quattro tipi di fluenza nativa (1) "the ability to fill time with talk" o "disc-jockey fluency," (2) "the ability to talk in coherent, reasoned and 'semantically dense' sentences," (3) "the ability to have appropriate things to say in a wide range of contexts," e (4) "the ability [...] to be creative and imaginative in language use". Raupach (1983) evidenzia come il parlato di parlanti nativi apparentemente corretti e fluenti sia comunque caratterizzato spesso da differenze nelle scelte sintattiche o false partenze ed enunciati incompleti. Questi aspetti del parlato non sempre dunque influenzano l'impressione generale di fluenza e spesso sono evidenti solo tramite un'osservazione attenta o una trascrizione dettagliata. Poiché questi fenomeni sono presenti anche nelle produzioni fluenti di parlanti nativi, sarà la distribuzione della loro frequenza piuttosto che la loro presenza o assenza l'elemento che offre informazioni. Davies (2003) sostiene che ciò che distingue le pause di un nativo da quelle di un non nativo è più facilmente la posizione piuttosto che la semplice durata. Nello studio condotto da Foster & Skehan (1996), ad esempio, nativi e non-nativi si differenziano proprio nella distribuzione delle pause, con i primi che le inseriscono prevalentemente in corrispondenza del confine di AS-unit. Risultati di questo tipo suggeriscono come un approccio che si limiti a un semplice conteggio delle pause possa avere dei limiti nella sua applicazione.

La discussione presentata in § 2.2.1.2 dimostra come le tre dimensioni CAF rilevino diversi aspetti della performance e come nella produzione esista in alcuni casi una correlazione tra le dimensioni CAF, in altri si generi un complesso sistema di trade-off. Quando l'obiettivo è individuare misure capaci di rilevare i livelli di proficiency e lo sviluppo longitudinale, è fondamentale conoscere non solo l'andamento delle singole misure prese individualmente, ma anche le loro interconnessioni. Peter Skehan e Peter Robinson nei loro modelli, seppure con le differenze che abbiamo già visto, identificano una tensione tra forma (complessità e accuratezza) e significato (fluenza). Skehan (1998) in particolare ha descritto queste tensioni sotto forma di ipotesi di trade-off, predicendo come l'attenzione a una dimensione abbia un'influenza sulle altre (cfr. § 2.2.1.2). Le tabelle 1-3 in Appendice riportano i principali studi a proposito. La ricerca sembra supportare l'idea di un trade-off tra complessità e accuratezza. Sebbene non sia semplice trarre generalizzazioni a partire dagli studi condotti fino ad oggi a causa delle differenze metodologiche. Così come evidenziato da Ellis (2005) nella sua presentazione di ricerche a proposito, i diversi tipi di planning favoriscono complessità o accuratezza, raramente entrambe. Analoghi i risultati relativi agli effetti della ripetizione del task o alle caratteristiche dei task, con uniche eccezioni gli studi di Gilabert (2007) e Tavokoli & Skehan (2005). La fluenza aumenta invece nella maggior parte degli studi, le differenze nelle misure impiegate non permettono però generalizzazioni in proposito.

Interessanti i risultati sui possibili trade-off tra dimensioni CAF ottenute da Wolfe-Quintero *et al.* (1998) nella loro meta-analisi. I ricercatori osservano infatti differenze tra misure di accuratezza da un lato e complessità e fluenza dall'altro nel modo in cui si relazionano ai livelli di proficiency. Le misure di accuratezza risultano in genere linearmente coerenti con giudizi olistici e cambiamenti sul breve periodo in classi intatte, mentre complessità e fluenza sono significativamente correlate con programmi, livello di scuola o livello di proficiency globale. Wolfe-Quintero *et al.* sulla base di questo dato concludono che le misure di fluenza e complessità si rifanno allo stesso costrutto, quello di

sviluppo, mentre l'accuratezza potrebbe essere legata ad un diverso costrutto, quello di errore. Lo studio qui condotto si pone l'obiettivo di approfondire queste questioni.

### 2.3. *L'approccio del testing*

I test linguistici sono sempre più diffusi in contesto scolastico e professionale, i risultati dei test vengono utilizzati per valutare le competenze di un individuo, ma soprattutto per prendere decisioni, non solo educative, su di lui. Ad esempio, un test linguistico potrebbe essere utilizzato per identificare gli studenti immigrati di una scuola o essere prerequisito per l'accesso ad un determinato percorso di studi o ancora per ottenere la cittadinanza. I test pertanto hanno la potenzialità di aiutarci a raccogliere informazioni utili a una varietà di individui, ma per realizzare appieno questa loro funzione, è necessario che il loro uso e la conseguente interpretazione dei risultati siano affidabili e validi (cfr. Bachman, 2004). In altre parole la comprensione del rapporto tra la performance di un candidato e le inferenze, predizioni o generalizzazioni fatte sulla base della performance stessa è il problema centrale del testing (*validity problem*). Una serie di variabili possono infatti influire sulla validità delle inferenze fatte a partire da un test: il task, il valutatore, la scala di proficiency, le caratteristiche del candidato e infine le possibili interazioni tra le variabili stesse (McNamara, 1996).

Se le produzioni di un apprendente variano, è essenziale determinare quale performance testare. Ecco che lo studio della variazione è un tema centrale anche nel testing, poiché può contribuire in modo sostanziale all'elaborazione di criteri efficaci per la selezione di task e di metodi di valutazione. I modelli di Bachman & Palmer (1996) e Skehan (1998) presentati in § 2.1, descrivendo i fattori che possono influire sulla performance determinando variazione, hanno stimolato nell'ambito del testing una varietà di studi che hanno cercato di approfondire il tema. Questa sezione presenta la

letteratura sulla variazione nel testing (§ 2.3.1) e sulla sua misurazione (§ 2.3.2).

### 2.3.1. La variazione

L'attenzione rivolta alla competenza comunicativa in campo acquisizionale e didattico ha avuto un impatto importante anche nella valutazione. A partire dagli anni '90 la letteratura sul testing si è infatti sempre più occupata del cosiddetto *language performance testing*, ossia la valutazione del candidato sulla base di performance elicitate, che hanno un certo grado di autenticità rispetto ai bisogni comunicativi del candidato stesso (per una presentazione dettagliata, vedi McNamara, 1997; Norris *et al.*, 1998). Se, come dimostra la letteratura acquisizionale, diversi task realizzano diversi gradi di accuratezza, fluenza, grammaticalità o comprensibilità, offrendo di conseguenza diverse prospettive sul grado di sviluppo dell'interlingua, quali strumenti possono permettere la creazione di un profilo affidabile e rappresentativo? Una soluzione riguarda l'uso di task più autentici, ossia più vicini agli usi reali della lingua. L'abilità linguistica così può essere misurata valutando la capacità dello studente di cavarsela in quella situazione. Chiaramente un approccio di questo tipo aumenterebbe la validità della prova, a scapito però di affidabilità e generalizzabilità (Hoekje & Linnel, 1994).

La ricerca ha tentato di risolvere questa tensione adottando contemporaneamente un approccio basato sui task (*task-centred approach*), per mantenere il legame con il mondo reale, e un approccio basato sul costrutto (*construct-centered approach*), per garantire comunque un certo grado di astrazione e generalizzazione. L'interesse per questa forma di testing ha determinato infatti il bisogno da un lato, di comprendere meglio la natura dei task utilizzati per elicitare la performance dall'altro, di sviluppare strumenti adeguati per valutarne l'utilità (Bachman, 2000). Proprio per la dimensione sociale che il testing assume, la valutazione è luogo di possibili conflitti politici

o scontro di valori, attraverso la validazione, chi si occupa di testing ha l'obbligo di considerare le possibili conseguenze di un test ed evitare quelle negative. Di qui l'attenzione nel testing per definizione del costrutto e validazione (Cronbach, 1971; Messick, 1989; Kane, 1992; Mislevy, Steinberg & Almond, 2003; per una presentazione dettagliata del concetto di validazione si vedano Bachman, 1990; e McNamara, 2006). Chi elabora un test deve seguire pertanto alcune tappe: innanzitutto deve specificare cosa vuole rilevare circa l'abilità del candidato sulla base di una performance; poi deve determinare che tipo di performance può supportare le osservazioni che si vogliono fare. Questo dipende chiaramente da una teoria che descrive le caratteristiche di una performance di successo, un costrutto che sarà chiaramente riflesso nelle scale di valutazione utilizzate; successivamente deve definire in termini generali i tipi di task in cui il candidato dovrà cimentarsi; e infine deve procedere alla valutazione, e conseguente elaborazione di un punteggio, sulla base della performance elicitata. Infine un aspetto fondamentale del processo sarà accompagnare il risultato di un test a dati che ne dimostrano la rilevanza rispetto alle osservazioni che vogliamo fare su quel test (*validity argument*). In altre parole è essenziale che un tester sia in grado di collegare una performance da un lato all'abilità linguistica del candidato, dall'altro all'uso di quell'abilità in un contesto reale. Per fare ciò sarà necessario capire: (a) come le variabili interne al candidato possano influenzare la performance; (b) come le variabili esterne al candidato possono influenzare la performance (per es. caratteristiche del task); e infine (c) come la performance in un task può essere costruita e interpretata in relazione a contesti d'uso reale (per es. utilizzando scale di proficiency o selezionando i task secondo criteri specifici). Diversi studi si sono dunque occupati del legame tra elementi quali task, valutatore, scala di proficiency e caratteristiche del candidato con la variabilità della performance, identificando una serie di fattori rilevanti. Di seguito si illustrano la letteratura in proposito.

Per quel che riguarda le caratteristiche dei candidati Kunnan (1993) riassume 20 studi sul tema, mostrando un impatto significativo sulla qualità



della performance di variabili quali livello di formazione, lingua madre, cultura o genere, professione (Hill, 1998), attitudine (Sasaki, 1996; Sparks *et al.*, 1998), conoscenze pregresse (Clapham, 1993; 1996) o caratteristiche personali (Berry, 1993).

Un'altra serie di studi si è occupata invece del processo di test-taking e delle strategie usate dai candidati nello svolgere un test. In questi studi si utilizzano spesso metodologie introspettive (per es. Buck, 1991; Wu, 1998) e interviste (per es. Purpura, 1993; 1997; 1999). In altri più innovativi si applicano le tecniche dell'analisi conversazionale o strumenti analitici di analisi dell'interlingua tipici di SLA per approfondire il legame tra intervista orale e tipologia di discorso (per es. van Lier, 1989; Berwick & Ross, 1996; Lazaraton, 1996) o caratteristiche linguistiche specifiche e livelli di proficiency (Banerjee *et al.*, 2007; Iwashita *et al.*, 2007; cfr. § 2.4)

Per quanto riguarda l'impatto del task un buon numero di studi ha mostrato che diversi task possono generare variazione nella performance e di conseguenza nella valutazione del test (cfr. Henning, 1983; Shohamy 1983; Shohamy, Reves & Bejarano, 1986). Così come Tarone (1983) ipotizzava che la variabilità non fosse un semplice fenomeno della performance, ma che piuttosto la conoscenza dell'apprendente possa essere variabile, anche nell'ambito del testing alcuni autori (Chalhoub-Deville, 1995a; 1995b) suggeriscono che la natura stessa del costrutto della L2 sia dipendente del contesto e pertanto vari da task a task. Chalhoub-Deville sostiene la sua proposta elaborando empiricamente task-specific dimensions di proficiency orale su tre diversi task, intervista, racconto di una storia e lettura ad alta voce. Anche nel caso in cui il task sia mantenuto costante, l'argomento può determinare variazioni nelle produzioni: Smith (1989) mostra come pronuncia, grammatica, fluenza e comprensibilità sono soggetti a variazione in relazione all'argomento, specifico di un campo disciplinare o generale, determinando differenze notevoli nei risultati del test. Nel suo studio infatti buona parte dei partecipanti non supera entrambi i test. Risultati analoghi sono stati ottenuti rispetto a fattori quali, concretezza, specificità dell'argomento, livello di

vocabolario, complessità sintattica, carico cognitivo, o quantità di processing richiesto (Alderson & Urquhart, 1985; Anderson *et al.*, 1991; Hale, 1988; Freedle & Kostin, 1993; Perkins *et al.*, 1995; Bachman *et al.*, 1996).

Un ulteriore filone di studi si è occupato dell'impatto che l'interlocutore ha sulle produzioni. Hoekje & Linnel (1994) valutano e confrontano tre strumenti per la valutazione orale, the SPEAK test, the OPI (*oral proficiency interview*) e IP test (*interaction performance*), che si differenziano per presenza o meno dell'interlocutore, numero degli interlocutori, potere dell'interlocutore rispetto a presa dei turni e gestione dell'argomento di conversazione, dimostrando come i tre strumenti mostrino differenze sostanziali riguardo all'uso linguistico considerato, in particolare rispetto a lunghezza dei turni e all'uso dei segnali discorsivi. Brown (2003; 2005) indaga l'influenza sulla performance di elementi quali stile interazionale e identità dell'intervistatore, mostrando come genere, identità professionale, esperienza pongano difficoltà diverse al candidato (per una presentazione dettagliata di queste ricerche vedi anche McNamara, Hill & May, 2002).

Anche rispetto a valutatore e strumenti di valutazione la letteratura evidenzia una serie di possibili cause di variazione. Diversi studi hanno mostrato l'impatto di variabili quali background del valutatore, eventuale conoscenza nativa della lingua target, livello di formazione (Brown, 1995; Chalhoub-Deville, 1996; McNamara, 1996; Weigle, 1994;1998; Schoonen *et al.*, 1997). In alcuni casi il processo di valutazione è stato studiato in dettaglio, così da identificare come il valutatore seleziona una certa valutazione. Linacre (1989) offre una discussione interessante sulla variazione associata ai valutatori e individua i principali punti di possibile divergenza: due valutatori possono differenziarsi per severità generale, oppure due valutatori possono essere più o meno severi rispetto a determinate strutture o compiti, o ancora due valutatori possono distinguersi nell'interpretazione di uno stesso criterio di valutazione, infine i valutatori possono distinguersi in coerenza. In genere gli studi condotti mostrano come, sebbene non sia né possibile e forse nemmeno auspicabile, eliminare del tutto le differenze, la formazione può per lo meno ridurle.

Diverse ricerche mostrano infatti come la formazione renda i valutatori più consistenti e riduca le differenze in severità (Lunz & Stahal, 1990; McIntyre, 1993; Weigle, 1994; Shohamy *et al.*, 1992).

Ricapitolando, una delle implicazioni della variazione sul testing riguarda il cosiddetto *validity problem*, ossia la standardizzazione del processo di valutazione stesso, con conseguenze in particolare per l'elaborazione di task e prove di valutazione. La questione è riassumibile nella seguente domanda: che tipo di elicitazione è meglio utilizzare, vista la sensibilità che l'interlingua dimostra verso minime differenze contestuali? E soprattutto, come misurarla? Questo studio attraverso l'esplorazione della variazione situazionale e longitudinale in diversi task si cercherà di offrire un contributo a queste questioni. La sezione che segue è dedicata alla presentazione della letteratura relativa alla misurazione della performance attraverso scale di proficiency.

### 2.3.2. La misurazione della variazione

Attualmente in Europa il punto di riferimento ufficiale per l'insegnamento e la valutazione degli apprendimenti linguistici è il *Quadro Comune Europeo di Riferimento per le lingue* (d'ora in poi QCER, Consiglio d'Europa 2001). Tale documento è il risultato del lavoro trentennale del Consiglio d'Europa nell'elaborazione di scale e descrittori dedicati alle competenze e agli usi linguistici con l'obiettivo di agevolare la cooperazione e il confronto nell'ambito dell'insegnamento/apprendimento delle lingue moderne, proponendo un linguaggio comune e favorendo la trasparenza e la coerenza in termini di definizioni e descrizioni di obiettivi di apprendimento, contenuti e metodi.

Sin dalla sua pubblicazione il QCER ha stimolato l'interesse di chi si occupa di insegnamento o valutazione, e la sua applicazione è stata oggetto di sperimentazione in diversi contesti di insegnamento e valutazione a livello europeo (cfr. Progetto Lingue Moderne; The Dutch CEF Construct Project,

Dialang, in Alderson, 2002; Alderson *et al.*, 2004; [www.coe.int/portfolio](http://www.coe.int/portfolio)) e in ambito italiano (cfr. Progetto Lingue 2000; Grego Bolli, 2006a; 2006b; [www.coe.int/portfolio](http://www.coe.int/portfolio)). Questi studi riconoscono le potenzialità e il valore del documento, e mettono in luce alcuni dei limiti nella sua applicazione al testing, evidenziando la necessità di ulteriori ricerche ai fini della sua implementazione. In questa sezione dopo una breve presentazione delle scale di proficiency, illustro le principali caratteristiche del QCER e le potenzialità e i limiti delle sue applicazioni al testing.

#### 2.3.2.1. Le scale di proficiency

In risposta ai bisogni di trasparenza dei sistemi educativi che danno sempre più valore alle descrizioni di cosa significhi esattamente essere a un determinato livello di proficiency, nell'ambito della valutazione si sono sempre più diffuse le scale di proficiency linguistica. Nella letteratura relativa al testing ci si riferisce spesso a tali scale con una varietà di termini “band scores, band scales, profile bands, proficiency levels, proficiency scales, proficiency ratings” (Alderson, 1991: 71) o “guidelines, standards, levels, yardsticks, stages, scales or grades” (De Jong, 1992: 43). I diversi nomi impiegati per indicare le scale di proficiency linguistica riflettono il fatto che le diverse scale prodotte abbiano origini diverse. Oltre a una differenziazione di origine, le scale possono essere poi ulteriormente distinte sulla base del loro uso. A tal proposito Alderson (1991: 72-4) introduce una loro classificazione funzionale distinguendo tre tipologie di scale: (a) *user-oriented*, con la funzione di riportare informazioni sul comportamento di una apprendente a un determinato livello; (b) *assessor-oriented*, con la funzione di guidare il processo di valutazione offrendo una descrizione della qualità della produzione; (c) *constructor-oriented*, con la funzione di guidare la costruzione di prove offrendo una descrizione dei task comunicativi che l'apprendente dovrebbe affrontare”.

L'elemento comune di questi strumenti è l'obiettivo che perseguono: offrire appunto una descrizione per livelli successivi della competenza linguistica. Tali scale possono pertanto contribuire alla trasparenza e alla coerenza nell'insegnamento linguistico ed essere utili per: (a) fornire uno "stereotipo" con cui l'apprendente può confrontare la sua posizione (Trim, 1978; Oscarson 1978; 1984); (b) fornire livelli che riflettono il curriculum scolastico e che possono essere formulati in termini di livello di proficiency raggiunta (Hargreaves, 1992); (c) aumentare l'affidabilità di valutazioni soggettive, in particolare rispetto alle competenze di produzione, fornendo uno standard comune (Alderson, 1991); (d) fornire linee guida per la costruzione di prove di valutazione (Alderson, 1991); (e) offrire punti di riferimento per chi deve gestire test d'ingresso, programmazioni, preparazione di materiale didattico, test di progresso e certificazioni finali (North, 1991); (f) permettere il confronto tra sistemi e gruppi di apprendenti utilizzando un metro comune (Bachman & Savignon, 1986; Carrol & West, 1989; cfr. North, 2000; 2002).

Diverse critiche sono state mosse alle scale di proficiency linguistica prodotte, in particolare rispetto al fatto che esse siano per la maggior parte scritte intuitivamente e quasi sempre manchino di una dimostrata validità (cfr. Brindley, 1998). Costruire una scala di proficiency non è certo un'operazione semplice. Si pongono infatti due tipi di problemi: uno descrittivo – le categorie utilizzate devono rifarsi ad un modello di proficiency – l'altro relativo alla misurazione: la scala di proficiency dovrà essere collegata a un modello di misurazione.

Dal punto di vista descrittivo, le scale di proficiency, in quanto composte da una serie di livelli ascendenti, sono strettamente collegate alle ricerche acquisizionali. I risultati ottenuti in letteratura non sono però ancora sufficientemente chiari per permettere una descrizione dettagliata dei percorsi di apprendimento: non c'è un'interpretazione univoca di competenza, proficiency e performance, né è stato elaborato un modello descrittivo condiviso della lingua in uso (cfr. § 2.1). Ne consegue, come evidenzia McNamara (1995), che spesso le scale tendono a offrire un modello

semplificato, per non dire impoverito, della lingua in uso, che con difficoltà può essere collegato a un modello teorico. Dal punto di vista della misurazione poi nella maggior parte dei casi sono scarsi o nulli i tentativi di dimostrare la validità della formulazione verbale dei descrittori prima della loro pubblicazione. Quando vengono fornite informazioni rispetto alla validità si tratta generalmente di analisi statistiche dell'affidabilità dei giudizi tra valutatori che utilizzano la scala dopo che essa è stata prodotta (cfr. McNamara, 1995). Il fatto che diversi utilizzatori possono interpretare coerentemente le scale di language proficiency non garantisce infatti che la scala sia una valida descrizione delle diverse tappe dell'acquisizione di una seconda lingua (Brindley, 1998: 22).

Per questa tensione esistente tra modelli teorici e modelli operativi sviluppati da esperti del settore e l'assenza di una descrizione empirica adeguatamente validata della complessità della proficiency linguistica, alcuni autori sono estremamente critici verso tali strumenti suggerendo come la costruzione di scale di proficiency o quadri di riferimento non sia operazione da intraprendere (cfr. Lantolf & Frawley, 1985; 1988; 1992). Come evidenzia North (2000: 13) nonostante i problemi connessi al descrivere un fenomeno complesso con poche parole e sulla base di una teoria incompleta, le scale di proficiency linguistica hanno la potenzialità di avere un'influenza positiva sull'orientamento, l'organizzazione e la valutazione dell'apprendimento linguistico e per questo meritano di essere sviluppate. Secondo l'autore infatti i limiti delle scale possono essere ridimensionati facendo una distinzione fondamentale tra quello che è un modello teorico di descrizione della natura della proficiency e quello che può essere un modello operativo utile a insegnanti e valutatori. Un modello operativo chiaramente è sempre una semplificazione di un modello teorico: fa riferimento alla teoria, ma ne reinterpreta gli elementi per renderli più accessibili in un particolare contesto. D'altronde, come suggeriscono anche altri autori, chi si occupa di valutazione degli apprendimenti linguistici non può certo attendere lo sviluppo di un modello di proficiency linguistica validato sperimentalmente prima di

elaborare criteri per la valutazione delle produzioni nella seconda lingua (Brindley, 1989: 56), ma piuttosto dovrà lavorare “con tassonomie che appaiono sensate anche se non completamente riconducibili a una descrizione teorica” (Hulstijn, 1985: 277).

L’arbitrarietà che dunque caratterizza gli standard di proficiency può essere limitata: (a) prendendo in considerazione teoria e ricerca; (b) assicurandosi che le decisioni prese si basino su un largo consenso per i contesti rilevanti; (c) tenendo conto della soggettività delle valutazioni individuali e di gruppo; (d) chiarendo esplicitamente quali sono i limiti della dichiarata validità empirica: anziché limitarsi ad affermare che una scala è valida, bisognerebbe specificare per cosa è valida (Henning, 1990: 379).

Il QCER tenta di superare i limiti discussi: le scale di descrittori non sono state formulate in maniera intuitiva, ma sono state sviluppate seguendo una metodologia rigorosa, effettuando analisi di tipo sia qualitativo che quantitativo (cfr. North, 2000). Dopo un’attenta revisione delle scale di proficiency esistenti rispetto alle categorie rilevanti per il modello teorico assunto nel QCER, gli autori hanno selezionato e prodotto una serie di descrittori. Questi sono stati prima sottoposti a una revisione con il coinvolgimento di insegnanti nel ruolo di informatori, poi calibrati utilizzando una versione scalare del *Rasch Model* (per una presentazione del modello vedi ad esempio McNamara, 1996, capitolo 6), ottenendo così una scala comune di descrittori della proficiency linguistica. Gli autori hanno dunque tentato di identificare categorie che potessero essere adeguate alla teoria, comprensibili a valutatori, insegnanti e studenti, e capaci di un’interpretazione sufficientemente condivisa in diversi contesti educativi (cfr. North, 2000). Nella sezione che segue il QCER viene presentato più in dettaglio.

### 2.3.2.2. Il Quadro Comune Europeo di Riferimento

Il QCER è riconosciuto come una delle pubblicazioni più influenti nel campo dell'apprendimento e della valutazione, è il prodotto finale di un progetto europeo iniziato negli anni '70 e coordinato dal Consiglio d'Europa (1988; 1992; 1996; 2002; 2003; 2005). In una prima fase è stata elaborata una serie di descrittori dei livelli di competenza, prima il *Threshold level* e poi i successivi *Waystage*, *Breakthrough Levels*, *Vantage*, *Effective* e infine *Mastery levels* (cfr. van Ek, 1975; 1977; Trim, 1978; van Ek & Trim, 1984; 1990; 1991; 1995). In una seconda fase si è mirato a riunire i diversi livelli in un unico quadro coerente (cfr. North, 2000). Tale quadro, il QCER appunto, sviluppato da John Trim, Brian North, Daniel Coste e Joseph Sheils è stato poi oggetto di un'intensa revisione a livello europeo da parte di esperti del settore. La complessità del lavoro ha portato alla pubblicazione di due versioni intermedie (Consiglio d'Europa, 1996; 1998), prima di arrivare alla versione finale del 2001 (Consiglio d'Europa, 2001). Inoltre lo sviluppo del QCER è stato documentato nei dettagli in diverse pubblicazioni (North & Schneider, 1998; North, 2000; Schneider & North, 2000) e una serie di studi di caso (Alderson, 2002; *Language Testing, Special Issue*, 2005). Per l'italiano si vedano Vedovelli (2002), Mezzadri (2004), Bosisio (2005), Grego Bolli, (2006a) e Jafrancesco, (2004).

Il QCER è un documento complesso e ambizioso, che nella propria dichiarazione di intenti al capitolo 1 si offre come strumento di lavoro a tutti coloro che si occupano di apprendimento linguistico: docenti e studenti naturalmente, ma anche autori di programmi e curricoli, responsabili delle politiche linguistiche, esponenti del mondo dell'editoria e enti certificatori. Gli autori del QCER descrivono così il documento e le sue finalità:

Il *Quadro Comune Europeo di Riferimento per le lingue* fornisce una base comune in tutta l'Europa per l'elaborazione di programmi, linee guida curriculari, esami, libri di testo per le lingue moderne ecc. Descrive in modo esaustivo ciò che chi studia una lingua deve imparare per usarla per comunicare e indica quali conoscenze e abilità deve sviluppare per agire in modo efficace. La descrizione riguarda anche il contesto culturale nel quale la lingua si situa.



Inoltre il *Quadro di Riferimento* definisce i livelli di competenza che permettono di misurare il progresso dell'apprendimento ad ogni stadio del percorso, nella prospettiva dell'educazione permanente. (QCER, p.1)

Il QCER descrive dunque la competenza linguistica attraverso una serie di scale accompagnate da introduzioni che forniscono utili informazioni su elementi rilevanti nell'insegnamento, nell'apprendimento e nella valutazione delle lingue. L'approccio di fondo che caratterizza il QCER è l'orientamento all'azione. La competenza, pertanto, viene formulata con la descrizione di cosa l'apprendente è in grado di fare nello svolgimento di un compito comunicativo in un determinato contesto utilizzando un certo repertorio di risorse linguistiche. Il documento permette di descrivere la competenza dell'apprendente secondo due dimensioni: una verticale quantitativa, data dalle scale globali, e una orizzontale qualitativa, costituita dai parametri per le attività comunicative e la competenza linguistico-comunicativa (QCER, capp. 4 e 5).

I descrittori, formulati in termini di 'dichiarazioni di capacità' (*Can-do statements*), sono raggruppati in tre livelli principali: A (livello elementare), B (livello intermedio) e C (livello avanzato). A sua volta ciascuno dei tre livelli è suddiviso in altri due sotto-livelli, dando origine a una scala a 6 bande distinte secondo le due modalità, dello scritto e del parlato, e secondo i tre parametri, della ricezione, della produzione e dell'interazione. Questi vengono poi dettagliati in descrittori applicabili ad attività ben definite quali *Intervistare ed essere intervistati* o *Prendere la parola in una discussione*. A questo si affianca poi una descrizione dei parametri ricezione, produzione e interazione secondo le categorie che catturano aspetti delle competenze di un apprendente, quali la *fluenza nel parlato*, l'*accuratezza grammaticale* e l'*appropriatezza sociolinguistica*. I descrittori così ottenuti sono stati selezionati, validati e ordinati lungo la scala (cfr. North & Schneider, 1998; North, 2000).

La procedura impiegata dai redattori del QCER per ordinare i descrittori sulla scala è estremamente raffinata, ma non esente da limiti. Per chi utilizza il QCER è importante capire il modello statistico impiegato, l'estensione del *Rasch Model* (cfr. Figueras *et al.*, 2005). Caratteristica fondamentale di tale

modello è l'assunzione dell'unidimensionalità del fenomeno. Questo significa che, secondo il modello, se il descrittore X sulla scala sta allo stesso punto del descrittore Y, allora un apprendente che riesce a fare X con un certo livello di successo riuscirà anche a fare Y. Questa intercambiabilità presupporrebbe che la proficiency nel parlato, il nostro descrittore Y ad esempio, potrebbe essere osservata attraverso un test di ascolto, il nostro descrittore X, e tale generalizzazione non sarebbe accettabile dal punto di vista teorico. Come riconoscono gli stessi autori del documento, un modello unidimensionale non può soddisfare pienamente le richieste di un sistema multidimensionale come quello rappresentato dalla competenza linguistica (cfr. North & Schneider, 1998; North 2000; cfr. § 2.1.2). Per questo il QCER stesso sottolinea in più punti di essere un documento di riferimento e mette in guardia su come un suo uso troppo rigido possa addirittura essere controproducente:

La costruzione di un quadro esaustivo, trasparente e coerente per l'apprendimento e l'insegnamento delle lingue non implica l'adozione di un sistema unico e uniforme. Al contrario, il QCER dovrebbe essere aperto e flessibile, in modo da poter essere applicato, con gli opportuni adattamenti, alle singole situazioni (QCER, p. 9).

Di conseguenza il QCER si definisce uno strumento multifunzionale, flessibile, aperto, dinamico, amichevole e soprattutto non dogmatico:

Gli esempi sono semplicemente indicativi, è possibile prenderne in considerazione alcuni, scartarne altri, e aggiungerne di nuovi. Il lettore deve sentirsi libero di farlo, dal momento che è suo compito definire gli obiettivi e i risultati (QCER, p. xiv).

Il lettore è invitato a usare le scale e i relativi descrittori in modo critico (QCER, p. xv).

È importante sottolineare come il QCER non è stato concepito per indicare come devono essere apprese, insegnate o valutate le lingue seconde, ma piuttosto per far emergere la consapevolezza, stimolare la riflessione e migliorare la comunicazione tra esperti del settore, ponendosi appunto come punto di riferimento. Come sottolineano gli stessi autori nell'introduzione, obiettivo del QCER è far scaturire domande e non dare risposte preconfezionate. In quanto 'di riferimento' il QCER intende promuovere riflessioni e discussioni e offrirsi come strumento per descrivere la diversità,

non certo indirizzare verso un sistema pan-europeo rigido e uniforme (Figueras *et al.*, 2005).

Il QCER pone un'attenzione particolare alla valutazione e alla certificazione ed esplicita chiaramente le sue possibili funzioni in questi due ambiti:

Le possibilità d'uso [rispetto alla valutazione della competenza linguistica] sono tre:

1. per specificare il contenuto di test ed esami
2. per definire i criteri in base ai quali un obiettivo di apprendimento si può considerare raggiunto, con riferimento sia alla valutazione di una particolare prestazione orale o scritta, sia alla valutazione continua dell'insegnante, dei pari e all'autovalutazione
3. per descrivere i livelli di competenza rilevati in test ed esami esistenti, in modo da fare confronti tra sistemi di certificazione diversi. (QCER, p. 24)

L'accettazione del QCER da parte di specialisti dell'insegnamento e della valutazione in diversi contesti europei ha motivato la realizzazione del Manuale *Relating language examinations to the Common European Framework of Reference for languages: learning, teaching, assessment* (Consiglio d'Europa, 2003, 2005; cfr. Figueras *et al.*, 2005). Obiettivo della pubblicazione è rispondere ai bisogni degli enti di valutazione che devono mettere in relazione i propri esami con i livelli del QCER, si presenta pertanto come una guida all'applicazione di buone prassi del testing e dei codici professionali del settore.

Poiché il QCER si pone come risorsa per specificare cosa valutare, come interpretare la performance e come confrontare diversi sistemi di certificazione, per riuscire a realizzare appieno queste sue funzioni deve soddisfare alcuni criteri: deve cioè essere esaustivo<sup>2</sup>, trasparente e coerente. Ecco come vengono definiti tali criteri nel QCER:

- **esaustivo**: dovrebbe cercare di specificare nel modo più completo possibile la gamma di conoscenze, capacità e usi linguistici (QCER p. 8)

---

<sup>2</sup> Nella versione inglese del QCER si usa l'aggettivo *comprehensive*: "It describes in a comprehensive way what language learners have to learn to do in order use a language" (CEFR, p.1). A questo proposito Figueras *et al.* (2005: 264) sottolineano che "The CEFR has aimed to be *comprehensive*, but this has to be understood in a very practical way, as not being synonymous with *exhaustive*". Mi sembra opportuno tenere conto di questa precisazione anche a proposito dell'italiano *esaustivo*, forse più cautamente sostituibile con *ampio*.

- **trasparente:** l'informazione deve essere formulata in modo chiaro ed esplicito, in modo da essere accessibile e facilmente comprensibile alle persone interessate (QCER p. 8)
- **coerente:** la descrizione deve essere scevra da contraddizioni interne (QCER p. 9)

Il QCER ha avuto un impatto positivo sulla valutazione in Europa e sicuramente è al momento l'elenco di descrittori di proficiency meno arbitrario. Nonostante i punti di forza però, l'uso del QCER attuale ai fini della valutazione pone alcuni problemi. Il QCER è certamente più esplicito e preciso di documenti che lo hanno preceduto, ma per sua natura rimane un sistema descrittivo astratto. Questo genera una discrepanza tra le descrizioni disponibili e i bisogni concreti di chi si occupa di valutazione. Bisogna chiedersi infatti se e in quale misura sia realmente possibile per gli operatori cui il QCER è destinato, applicare concretamente il documento o parti di esso per il proprio lavoro di tutti i giorni, poiché il QCER non è né di facile interpretazione, né di facile utilizzo, ma richiede piuttosto che gli operatori siano accompagnati nella sua applicazione. Due sono i principali punti critici da tenere in considerazione: interpretazione e quantificazione (*linking problem*):

Per quel che riguarda l'interpretazione il problema è generato dal fatto che i *Can-do statements* non sempre coincidono con comportamenti direttamente osservabili. Essi devono essere intesi come categorie astratte e richiedono pertanto interpretazione da parte di chi li usa per costruire prove o valutare la performance. È necessario chiedersi quindi fino a che punto il QCER può essere uno strumento utile per la valutazione di diverse attività linguistiche e competenze. I problemi riscontrati nell'uso delle scale sembrano riguardare in particolare quattro aspetti: terminologia, lacune nelle specificazioni necessarie per il testing, incoerenze nei descrittori e assenza di definizione per alcune strutture. Vediamo un esempio a titolo illustrativo. Il descrittore relativo a *coerenza-coesione* per il livello A2+ è così formulato:

[l'apprendente] è in grado di collegare frasi semplici usando i connettivi più usuali per raccontare una storia o descrivere qualcosa, realizzando un semplice elenco di punti.

Quello del livello B1 recita invece:

è in grado di collegare una serie di elementi relativamente brevi e semplici in una sequenza lineare per punti.

È evidente che un valutatore rischia di trovarsi in seria difficoltà nel momento in cui deve determinare se l'apprendente che sta valutando dal punto di vista della coerenza-coesione si trova a livello A2+ o a livello B1: come definire le *frasi semplici*? E come distinguerle dagli *elementi relativamente brevi e semplici*? E quali sono i *connettivi più usuali*? Curiosamente, il descrittore del livello precedente (A2) risulta invece più preciso nelle indicazioni, giacché specifica a quali indicatori linguistici fa riferimento: *È in grado di collegare gruppi di parole con connettivi semplici quali “e”, “ma” e “perché”*. Perché tale esemplificazione è presente solo per questo livello e non per i due successivi? Casi di questo tipo sono tutt'altro che isolati nel QCER (cfr. Alderson *et al.*, 2004; Ferrari & Nuzzo, in stampa). Per rendere il QCER adatto ai bisogni del testing sarà necessario, da parte della ricerca, approfondire meglio in futuro il legame tra i diversi livelli e la varietà del repertorio di strutture linguistiche e forme d'uso. Diversi autori sottolineano a questo proposito l'utilità ad esempio dell'analisi di corpora (cfr. Hawkey & Barker, 2004) o il raffronto con le sequenze di acquisizione individuate nelle diverse lingue (cfr. Brindley, 1998). Nel tentativo di ridimensionare i limiti sopra descritti il Manuale sottolinea in più parti l'importanza della formazione dei valutatori e della verifica della validità e coerenza delle loro interpretazioni e auspica che la valutazione degli apprendimenti linguistici sulla base del QCER diventi parte integrante del curriculum degli insegnanti: una buona conoscenza del QCER è prerequisito essenziale per il raggiungimento di un maggior grado di affidabilità nei giudizi. La formazione specifica degli insegnanti nell'uso di questo strumento dovrebbe prevedere sia una fase iniziale di familiarizzazione con il QCER sia dei momenti di esercitazione e di confronto tra valutatori (*benchmarking*), con l'obiettivo di uniformare il più possibile le interpretazioni dei descrittori e delle performance. A questo proposito il Consiglio d'Europa

sta mettendo a punto ulteriori materiali di supporto: una banca dati europea di esempi di performance e di item rappresentativi (cfr. [www.coe.int/portfolio](http://www.coe.int/portfolio)).

Per quel che riguarda il problema di quantificazione tre sono gli aspetti da considerare: la quantificazione dei *Can-do statements*, la quantificazione della performance, la messa in relazione delle due quantificazioni. Qual è la procedura con cui una persona può essere collocata in uno dei livelli descritti sulla base della performance in modo che il livello assegnato corrisponda effettivamente al livello descritto nel QCER? I *Can-do statements* sono caratterizzati da una dimensione qualitativa e una quantitativa. Se nell'applicazione dovessimo limitarci alla dimensione qualitativa, il risultato della valutazione sarebbe un profilo qualitativo fatto di un elenco di *Can-do statements*. Un approccio di questo tipo sarebbe sicuramente appropriato da un punto di vista teorico e psicometrico, ma poco economico da un punto di vista pratico. Diventa quindi necessario rifarsi all'aspetto quantitativo del QCER che permette di ordinare le celle di descrittori lungo un continuum. Chi utilizza le griglie avverte la necessità di sapere quale sia il livello di qualità della performance che ci si aspetta dagli apprendenti nello svolgere un determinato compito a un determinato livello. Occorrerebbe cioè che fossero chiaramente esplicitati i criteri di valutazione di *come* un apprendente sa fare qualcosa. Tali criteri dovrebbero essere strettamente connessi con i parametri contestuali e teorici che influenzano l'abilità che si sta misurando. In altre parole per applicare le griglie in modo efficace sarà necessario rispondere ad almeno due domande: Qual è la percentuale di item a un determinato livello che l'apprendente deve superare, e/o qual è la proporzione di descrittore che un apprendente deve soddisfare per essere assegnato a un determinato livello? Quanti campioni scritti o orali sono necessari per assegnare un apprendente a un particolare livello?

Sulla base di osservazioni di questo tipo, diversi autori che l'hanno esaminato criticamente, come ad esempio Weir (2005: 281-282), sostengono che il QCER nella sua forma attuale non è sufficientemente ampio, coerente e trasparente per un uso efficace nel testing linguistico. Poiché il quadro non è

stato realizzato per rispondere a tutte le esigenze del testing, “it will require considerable, long-term research, much reflective test development by providers, and prolonged critical interaction between stakeholders in the field to address these deficiencies.” (Weir, 2005: 283; cfr. Alderson, 2004:1). In questo studio pertanto si cercherà di avere un approccio sperimentale all’uso del QCER, come del resto fortemente auspicato anche dal Consiglio d’Europa stesso.

#### 2.4. *L’interfaccia acquisizione-testing*

I due campi della linguistica acquisizionale e del testing hanno molto in comune, tuttavia diversi elementi li mantengono piuttosto distinti: il focus, la metodologia di analisi e infine i vincoli entro i quali operare. Rispetto al focus, se le ricerche di tipo acquisizionale tendono a concentrarsi su dettagli o aspetti specifici della performance con l’obiettivo di descrivere il processo di apprendimento linguistico e individuare gli elementi in grado di favorirlo, le ricerche relative al testing cercano di assumere uno sguardo più globale e ottenere gli strumenti per misurare la proficiency in modo affidabile e generalizzabile a contesti reali, sulla base di una performance elicitata; rispetto alla metodologia di analisi, se le ricerche acquisizionali hanno sviluppato strumenti descrittivi analitici e protocolli etnografici o sperimentali per l’analisi dell’interlingua, le ricerche sul testing utilizzano principalmente procedure statistiche da applicare ai risultati di un test; e infine rispetto ai vincoli, nonostante SLA offra indicazioni utili all’insegnamento, l’applicazione pratica dei risultati non è tra gli obiettivi primari, il testing invece è più orientato alla soluzione di richieste concrete, dovendo arrivare alla produzione di test.

Diversi studiosi hanno però riconosciuto come tali differenze possano costituire il punto di partenza per sviluppare interazioni fruttuose, capaci di avvantaggiare entrambi gli ambiti. In questa sezione si presenterà la letteratura relativa all’esplorazione delle differenze e dei punti di contatto tra acquisizione

e testing (§ 2.4.1), poi si illustreranno i vantaggi di una maggiore scambio tra i due campi in particolare rispetto al tema oggetto di questo studio, ossia la descrizione e misurazione della proficiency linguistica. Si farà riferimento ai temi della validità dei costrutti, dell'affidabilità degli strumenti di misurazione e infine dello sviluppo di scale di proficiency (§ 2.4.2).

#### 2.4.1. L'acquisizione e il testing

La natura della relazione tra acquisizione e testing viene discussa per la prima volta in modo specifico in un testo del 1998, dal titolo *Interfaces Between Second Language Acquisition and Language Testing Research*, curato da Lyle F. Bachman e Andrew D. Cohen. I vari contributi affrontano il tema da prospettive diverse: Carol A. Chapelle discute della definizione dei costrutti e del tema della loro validità nelle ricerche acquisizionali, Elaine Tarone esamina le ricerche sulla variazione interlinguistica e evidenzia le possibili implicazioni per il testing, Geoff Brindley si concentra sulla relazione tra scale di valutazione e ricerca acquisizionale, Dan Douglas discute il tema del metodo di valutazione nella ricerca acquisizionale di tipo context-based, Elana Shohamy usa il caso dell'analisi del discorso per evidenziare come i due campi non tengano conto dei rispettivi risultati e infine viene ripubblicato un contributo di Lyle F. Bachman del 1988 che riassume in prospettiva generale le possibili interazioni tra i due campi. La pubblicazione, evidenziandone i diversi potenziali benefici, sostiene l'importanza di una maggiore interazione tra i due ambiti. Questa prima discussione è stata seguita da altre pubblicazioni (Shohamy, 1998; 2000; Ellis, 2001; Douglas, 2001; Laufer, 2001; Hulstijn & Schoonen, 2006; Read, 2007) e ha portato recentemente alla costituzione di un gruppo di ricerca europea sul tema (SLATE, con il coordinamento di Jan Hulstijn). Vediamo ora in dettaglio le principali differenze e somiglianze tra i due ambiti. La discussione che segue si basa in buona parte su Bachman & Cohen (1998b) e Shohamy (1998).



E' possibile identificare almeno cinque macro-differenze tra acquisizione (SLA) e testing (LT). La prima è di tipo strutturale: ricercatori di SLA e ricercatori di LT appartengono a comunità distinte, ciascuna con i suoi convegni e le sue pubblicazioni specialistiche. Pochi i ricercatori che sono membri attivi di entrambe le comunità.

La seconda riguarda invece l'orientamento generale. SLA è più legata alla teoria, in quanto mira alla costruzione o validazione di teorie in grado di spiegare il processo di apprendimento. Pertanto, nonostante la disciplina sia collegata all'insegnamento, raramente l'obiettivo è elaborare proposte didattiche o ottenere applicazioni immediate dei risultati. Al contrario il testing ha un orientamento più applicato: la maggior parte degli studi LT sono infatti condotti entro progetti legati a test di proficiency e mirano a dimostrare che il sistema di valutazione impiegato è valido, funzionale e affidabile. La teoria in questo caso ha l'obiettivo di mettere in relazione la performance in un test con quella reale.

La terza differenza può essere riassunta dalla distinzione tra competenza e proficiency. In qualche modo SLA è più orientata alla descrizione della rappresentazione mentale, ossia della competenza e dei processi a essa collegati, mentre LT è orientato alla descrizione dell'abilità a usare tale conoscenza linguistica.

La quarta differenza riguarda gli strumenti di misurazione. I ricercatori di SLA hanno messo a punto sistemi di analisi dell'interlingua legati al conteggio delle occorrenze di determinate strutture in contesti obbligatori o misure analitiche più globali, come quelle di complessità, accuratezza e fluenza, che offrono descrizioni estremamente dettagliate della performance, a costo però di una loro applicazione piuttosto dispendiosa in termini di tempo. I ricercatori LT dovendo sviluppare o validare strumenti di misurazione da utilizzare con popolazioni numerose valutano la performance a partire da giudizi olistici, solitamente realizzati da valutatori esperti a partire da scale di proficiency. In altre parole, anche se i task di elicitazione possono essere simili, le misure applicate si differenziano in modo sostanziale.

Infine la quinta differenza riguarda aspetti più tecnici della ricerca. Gli studi SLA solitamente coinvolgono un numero limitato di partecipanti su base volontaria, mentre le ricerche LT si basano su test che hanno solitamente un impatto rilevante sulla vita dei candidati. Di conseguenza, se i ricercatori SLA possono elaborare con una certa flessibilità i loro task di elicitazione, i ricercatori LT sono più spesso legati a vincoli imposti dagli *stake-holders* che richiedono un certo tipo di valutazione.

Le differenze appena descritte sono bilanciate da una serie di possibili interazioni. Shohamy (1998) ne identifica sei: tre rappresentano possibili contributi di LT a SLA e tre di SLA a LT.

Le aree in cui LT può contribuire a SLA sono: (a) definizione del costrutto abilità linguistica. Una componente importante delle ricerche sul testing riguarda la definizione dell'abilità linguistica, LT può offrire modelli utili per verificare la validità dei risultati; (b) applicazione di risultati LT per testare ipotesi SLA. I ricercatori LT ottengono diversi tipi di dati circa il comportamento linguistico durante un test, analizzando ad esempio come specifiche strutture possano variare nelle condizioni di testing; (c) scelta di criteri validi per la selezione di test e task. Un importante contributo di LT a SLA riguarda lo sviluppo di criteri per sviluppare task in grado di elicitare dati acquisizionali validi e affidabili.

Le aree in cui SLA può contribuire a LT sono: (a) identificazione delle componenti linguistiche da elicitare e dei criteri per la loro misurazione. La ricerca SLA può identificare i costrutti necessari a chi si occupa di testing e offrire pertanto indicazioni utili per la costruzione di test. Ad esempio, le ricerche su variazione e task danno informazioni importanti sul ruolo di variabili quali interlocutore, argomento o status sociale sulla produzione, con una ricaduta importante sulla selezione dei task. Analogamente, rispetto alla definizione dei costrutti da misurare, un tester può basarsi sugli studi SLA; (b) costruzione di task per la valutazione. I ricercatori SLA hanno utilizzato una varietà di strumenti di elicitazione. Uno dei possibili contributi di SLA a LT potrebbe riguardare proprio l'ampliamento di tale repertorio; e infine (c)

informare i tester sui fattori che influenzano l'apprendimento. I ricercatori SLA possono suggerire ai tester possibili aree problematiche. Ad esempio, se gli studi SLA concludessero che la L1 ha un'influenza importante, allora i tester potrebbero adattare i test al background linguistico dei candidati.

Nonostante l'utilità reciproca di una proficua collaborazione, a tutt'oggi le interazioni tra i due campi sono piuttosto limitate. Shohamy (2000) basandosi sull'analisi di due diverse fonti di dati, ossia articoli apparsi in pubblicazioni scientifiche recenti e interviste con studiosi SLA, mette in luce proprio come l'interazione tra acquisizione e testing sia nella pratica ancora limitata. Di qui, il tentativo, con questo studio, di considerare entrambi gli approcci nello studio della variazione. In particolare ci interessano tre possibili aree di contatto: validità dei costrutti, qualità degli strumenti di misurazione, e infine sviluppo di scale di proficiency. Nella sezione che seguo illustro più in dettaglio questi temi.

#### 2.4.2. I costrutti, la misurazione e le scale di proficiency

Nelle ricerche acquisizionali gli strumenti di misurazione quantitativa e i test linguistici o i task di elicitazione hanno un ruolo piuttosto rilevante, fornendo le basi per la raccolta e la codifica dei dati. Ciononostante, come evidenziano Norris & Ortega (2003) nella loro presentazione della letteratura sulla misurazione in SLA, uno dei limiti principali riguarda proprio l'attenzione alle possibili fonti di errore di misurazione di tali strumenti. Questo è testimoniato dall'assenza nella maggior parte delle ricerche di dati relativi all'affidabilità degli strumenti impiegati. Questa lacuna sarebbe difficilmente comprensibile nella prospettiva LT, che controlla con attenzione l'affidabilità degli strumenti attraverso specifiche procedure statistiche.

L'affidabilità comunque è solo un aspetto di una buona misurazione. Di importanza ancora maggiore è la corrispondenza tra base teorica e dato empirico, ossia il ricercatore deve poter specificare cosa una data misura rileva.

Il concetto di validità del costrutto ha un ruolo centrale in LT (cfr. Cronbach & Meehl, 1955; Messick, 1989; Bachman & Palmer, 1996) e riguarda proprio l'interazione tra motivazione teorica (costrutto) e evidenza empirica (punteggio nel test o altro dato). Messick (1996) mette in luce due possibili rischi: sottorappresentazione del costrutto (*construct underrepresentation*) e varianza irrilevante al costrutto (*construct-irrelevant variance*). La sottorappresentazione riguarda i casi in cui la misura applicata non richiede al candidato di utilizzare tutta la gamma di abilità o conoscenze necessarie per svolgere il task nella vita reale. La varianza irrilevante al costrutto riguarda invece i casi in cui la misura è influenzata da fattori che non hanno nulla a che vedere con ciò che voglio misurare.

Nel caso di SLA vi è rischio di sottorappresentazione dei costrutti in almeno due casi. Il primo è legato all'operativizzazione: i diversi approcci teorici SLA – generativo, interazionista e emergentista – hanno identificato costrutti rilevanti, ma senza offrire né definizione del tipo di dato empirico che può sostenere una certa interpretazione, né validazioni sperimentali (cfr. Pallotti, 2007). Il secondo caso è legato alla metodologia e all'uso di protocolli cross-sectional, che portano inevitabilmente a generalizzazioni, piuttosto che protocolli longitudinali, più affidabili nella descrizione del processo di sviluppo. Ciò è ancora più evidente per i molti casi in cui i livelli dei diversi gruppi di apprendenti da confrontare sono stabiliti in modo intuitivo o attraverso strumenti non adeguati (cfr. Norris & Ortega, 2003).

In entrambi i casi SLA può rafforzare la propria analisi e l'affidabilità dei risultati appoggiandosi alla letteratura e alle consolidate procedure di LT. Vediamo ora un'area in cui è LT che può ricavare elementi utili da SLA: la creazione di scale di proficiency.

Il *proficiency movement*, richiedendo l'osservazione degli apprendenti in diversi contesti comunicativi, ha reso necessario lo sviluppo di scale di valutazione in grado di descrivere i diversi livelli di performance e assistere dunque i valutatori nell'applicazione coerente dei criteri di valutazione. Abbiamo già presentato in § 2.3.2 le scale di proficiency e il QCER e discusso

di alcuni limiti che li caratterizzano. Le scale di proficiency nella maggior parte dei casi non si appoggiano a basi teoriche e i costrutti che esse descrivono non sono né definiti esplicitamente né validati. Quando una forma di validazione è stata applicata, come per il QCER, il punto di riferimento è dato dai giudizi degli insegnanti (cfr. North & Schneider, 1998). Inoltre la formulazione dei descrittori in termini di *Can-do statements* non fanno riferimento a strutture linguistiche specifiche associate a quel determinato livello (cfr. § 2.3.2.2). Le ricerche di SLA potrebbero certo contribuire a colmare queste due lacune.

Esempi in questo senso sono rappresentati da alcuni studi che hanno tentato di esplorare le possibili interazioni tra i due campi. Brindley (1998) ad esempio, riprendendo la proposta di Pienemann & Mackey (1992) suggerisce la possibilità di applicare la Teoria della Processabilità alla valutazione della proficiency orale. Banerjee *et al.* (2007) hanno studiato la relazione tra bande di livello IELTS e features linguistiche nella produzione scritta, misurando quindi dimensioni della performance scritta quali uso di strumenti di coesione, ricchezza lessicale, complessità sintattica e accuratezza grammaticale a diversi livelli, e constatando come tutte le dimensioni (a eccezione della complessità sintattica) sono informative rispetto ai progressi in proficiency. In uno studio analogo Iwashita *et al.* (2007) confrontano la relazione tra misure analitiche di accuratezza, complessità, vocabolario, pronuncia e fluenza e giudizi olistici in cinque diversi task del TOEFL iBT e a cinque diversi livelli di proficiency. Anche in questo caso le categorie selezionate sono significative nella distinzione tra livelli. Studi di questo tipo non possono che contribuire alla descrizione e misurazione della proficiency linguistica, mostrando l'utilità dell'applicazione in un contesto LT di misure sviluppate in ambito SLA.

Ricapitolando, chi si occupa di SLA può imparare dal testing l'importanza di una verifica di validità e affidabilità degli strumenti di elicitazione e misurazione, oltre che di una più adeguata definizione operativa dei costrutti teorici e ritrovare quindi nella letteratura LT procedure e modelli operativi da utilizzare a questi scopi. Chi si occupa di LT può migliorare i propri strumenti di descrizione della performance inglobando descrizioni

dettagliate di aspetti linguistici caratteristici dei diversi livelli di proficiency. In questo studio si è cercato di cogliere questi stimoli e prestare attenzione a questi aspetti. Da un lato la metodologia di ricerca include sia discussioni circa il legame tra costrutti e relative misure sia controlli sperimentali circa l'affidabilità dell'applicazione delle misure stesse. Dall'altro, attraverso una delle domande di ricerca, che riguarda proprio il confronto tra misurazioni CAF e valutazioni QCER, si cercherà di ricavare implicazioni utili per una formulazione più dettagliata di descrittori di profili di proficiency.

## 2.5. *I task*

Con il diffondersi dell'approccio comunicativo i task e la loro relazione con performance e acquisizione hanno ricevuto una crescente attenzione in ambito didattico, ma anche in campo acquisizionale e nel testing. La vitalità di tali ricerche è testimoniata da un numero di pubblicazioni legate appunto a task per la didattica, la ricerca e il testing (Bygate *et al.*, 2001; Ellis, 2003; Long & Crookes, 1992; Skehan, 2003). Chiaramente il tipo di approccio ai task sviluppato in ciascuno dei tre ambiti è strettamente legato al contesto d'uso: la linguistica acquisizionale pertanto esplora i task come strumento per indagare il processo di apprendimento, la didattica come possibile attività per lo sviluppo linguistico, e infine il testing come modalità di valutazione delle competenze. In questa sezione si offrirà una definizione di task (§ 2.5.1) e si discuterà più in dettaglio dell'uso del task nei tre ambiti: didattica, ricerca e testing (§ 2.5.2).

### 2.5.1. La definizione di task

I task sono interpretati e definiti in modo diverso dai vari gruppi di persone che li utilizzano. Lo scopo di questa sezione è illustrarne le diverse interpretazioni e definizioni per poi selezionare quella utile per questo studio.

I task sono utilizzati in almeno tre contesti: didattica, ricerca e testing. Nel primo caso l'enfasi è rivolta all'efficacia nell'uso in classe, nel secondo invece è posta su ciò che accade all'apprendente durante lo svolgimento di un task (o di una serie di task) o sui cambiamenti che avvengono nell'interlingua e infine l'ultimo indaga la natura della performance così elicitata. La tabella 5 proposta da Bygate *et al.* (2001) illustra i tre possibili contesti d'uso dei task e mostra come in tutti i casi l'interesse possa essere di tipo sia pedagogico che sperimentale.

**Tabella 5. Lo studio dei task,  
da Bygate, Skehan & Swain, 2001**

	<b>pragmatic/pedagogic</b>	<b>research</b>
<b>teachers and teaching</b>	<ul style="list-style-type: none"> <li>- task as a unit of work in a scheme of work</li> <li>- interlinked activity sequences</li> <li>- developing thematic unit</li> <li>- methods of involving learners</li> <li>- deliberate starting point for unknown direction or explorations</li> </ul>	<ul style="list-style-type: none"> <li>- task as researchable unit</li> <li>- neat, cross sectional approach</li> <li>- relatively brief time interventions</li> <li>- focus on the isolation of variables</li> <li>- search for 'effects' through manipulation</li> </ul>
<b>learners and learning</b>	<ul style="list-style-type: none"> <li>- learners orientation and autonomy</li> <li>- task reinterpretability</li> <li>- interactive development through collaboration of groups of learners</li> <li>- authenticity of response</li> </ul>	<ul style="list-style-type: none"> <li>- extent to which learning processes are catalysed</li> <li>- identification of theorised methods of operationalising constructs and measuring dependent variables</li> <li>- research design to probe: <ul style="list-style-type: none"> <li>- salient task variables</li> <li>- salient task conditions</li> </ul> </li> </ul>
<b>testing</b>	<ul style="list-style-type: none"> <li>- formative evaluation</li> <li>- provision of structured feedback on communication</li> <li>- reactive, unstandardised and individual based</li> </ul>	<ul style="list-style-type: none"> <li>- summative evaluation</li> <li>- task as a format</li> <li>- comparability and standardisation</li> <li>- issues in performance assessment</li> </ul>

Sulla base del contesto in cui viene affrontata – insegnamento, ricerca o testing – la definizione di task può essere elaborata così da mettere in luce aspetti diversi. In ambito didattico i task possono essere intesi sia come attività contenute (Nunan, 1989) o sequenze più lunghe e composite che occupano diverse lezioni (ad es. il *task cycle* proposto da Willis, 1996) o addirittura temi da cui partire per sviluppare una serie di attività (ad es. *project work*). In ogni caso l'obiettivo è quello di produrre attività pedagogiche efficaci. Nel caso della ricerca acquisizionale invece il task è visto come attività che mira a elicitarre dati utili per il ricercatore, l'interesse è sempre motivato da domande di tipo teorico. Di conseguenza di estrema rilevanza saranno aspetti quali standardizzazione e validità ecologica. In un approccio quantitativo si cercherà di ottenere valutazioni statistiche circa la robustezza dei dati raccolti e in un approccio più qualitativo si condurranno analisi a partire dalle trascrizioni delle interazioni. In entrambi i casi l'obiettivo è proporre metodi d'indagine che permettano lo studio del processo di apprendimento in modo affidabile e replicabile. Infine nell'ambito del testing l'attenzione è rivolta a come rendere il task uno strumento valido, affidabile e praticabile ai fini della valutazione della proficiency. L'obiettivo qui è dunque l'esplorazione di come funzionano i task e quali influenze sistematiche hanno sulla produzione linguistica.

Nei tre diversi contesti sono state proposte varie definizioni di task, le principali sono riportate qui sotto. Alcune di queste sono più ampie ed inclusive, altre definiscono il task secondo criteri più restrittivi. Un elemento di distinzione riguarda la definizione del legame tra task e comunicazione autentica: in alcuni casi l'autenticità è definita attraverso la corrispondenza con attività quotidiane (per es. Long, 1985a), in altri è data dalla natura della risposta dell'apprendente e dalla sua somiglianza, diretta o indiretta, con la lingua usata in contesti di comunicazione autentica (per es. Skehan, 1998).

A task is a piece of work undertaken for oneself or for others, freely or for some reward. Thus examples of tasks include painting a fence, dressing a child ... In other words, by 'task' is meant the hundred and one things people *do* in everyday life, at work, at play, and in between. (Long, 1985a)



A piece of work or an activity, usually with a specific objective, undertaken as part of an educational course, at work, or used to elicit data for research. (Crookes, 1986)

One of a set of differentiated, sequencable, problem-posing activities involving learners and teachers in some joint selection from a range of varied cognitive and communicative procedures applied to existing and new knowledge in the collective exploration and pursuance of foreseen or emergent goals within a social milieu. (Candlin, 1987)

An activity which required learners to arrive at an outcome from given information through some process of thought and which allowed teachers to control and regulate that process was regarded as 'task'. (Prabhu, 1987)

Any structured language learning endeavour which has a particular objective, appropriate content, a specified working procedure, and a range of outcomes for those who undertake the task. 'Task' is therefore assumed to refer to a range of workplans which have the overall purpose of facilitating language learning – from the simple and brief exercise type, to more complex and lengthy activities such as group problem-solving or simulations and decision-making. (Breen, 1987)

A piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than on form. (Nunan, 1989)

A task [is] any activity in which a person engages, given an appropriate setting, in order to achieve a specifiable class of objectives. (Carroll, 1993)

We define a *language use task* as an activity that involves individuals in using language for the purpose of achieving a particular goal or objective in a particular situation. (Bachman & Palmer, 1996)

Tasks are always activities where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome. (Willis, 1996)

A task is an activity in which

- meaning is primary
- learners are not given other people's meanings to regurgitate
- there is some sort of relationship to comparable real-world activities
- task completion has some priority
- the assessment of the task is in term of the outcome. (Skehan, 1998)

Selezionare una definizione appropriata di task, significa rifarsi all'uso per cui esso è inteso. Poiché in questo studio l'obiettivo è misurare la variabilità della performance, ci si riferirà al task riprendendo la seguente definizione proposta da Bygate *et al.* (2001:12).

A task is a contextualised, standardised activity, relatable to learner choice or to learning processes, which requires learners to use language, with emphasis on meaning, to attain an objective, and which elicits data which may be the basis for research.

In questo studio il task è usato come strumento di elicitazione in grado di favorire un'osservazione replicabile ed ecologicamente valida del processo di apprendimento. L'analisi dei dati mira però ad offrire elementi di riflessione utili anche a chi utilizza i task per scopi di insegnamento o testing. Nella sezione che segue pertanto si descriverà in dettaglio la letteratura relativa ai tre contesti d'uso dei task.

#### 2.5.2. I task per insegnamento, SLA e testing

L'approccio comunicativo è una sorta di *post-method approach* all'insegnamento (Brumfit & Johnson, 1979; Brumfit, 1984; 1988), che favorisce l'uso di una varietà di tecniche, piuttosto che offrire all'insegnante pacchetti preconfezionati. Riprendendo idee sviluppate in ambito pragmatico, sociolinguistico o dell'analisi del discorso, pone al centro dell'attenzione l'importanza di un uso contestualizzato e autentico della lingua nella classe di L2 (Brumfit & Johnson, 1979). A partire dagli anni '70 si sono sviluppate dunque diverse proposte didattiche in questo senso. Le prime interpretazioni hanno portato alla creazione di attività in grado di promuovere un uso più autentico della lingua, quali *information gap activities* (Harmer, 1991) o *jigsaw activities* (Geddes & Sturtridge, 1978; per una discussione più recente vedi Wesche & Skehan, 2002). Inizialmente ci si riferiva a queste attività con il termine generico 'attività comunicative', poi sempre più spesso si è utilizzata la parola task. Questi sviluppi pedagogici hanno influenzato notevolmente la glottodidattica, portando sia all'elaborazione di diverse proposte di approcci didattici o sillabi basati interamente sul task (Long & Robinson, 1998; Nunan, 1989; Prabhu, 1997; Willis & Willis, 1988), sia di procedure di valutazione (Morrow, 1979). Contemporaneamente la ricerca acquisizionale si è interessata al task come strumento di elicitazione, oltre che costrutto pedagogico da indagare.

Un approccio all'insegnamento basato sul task parte dal presupposto che una lingua si apprende per comunicare, pertanto poiché il significato è primario rispetto alla forma, le attività devono essere in grado di favorire un'interazione più autentica. Il punto di criticità di questo approccio però è che attirando l'attenzione verso il significato, si rischia di distrarre l'apprendente dalla forma. La sfida dunque è riuscire a stimolare una quantità sufficiente di *focus on form* così da permettere lo sviluppo del sistema interlinguistico senza compromettere la naturalezza dello scambio comunicativo generato dal task. Ottenere ciò significa sviluppare gli strumenti per comprendere meglio almeno tre elementi: l'impatto del task design e delle condizioni di somministrazione sulla performance, l'impatto della selezione dei task sull'uso linguistico e infine le relazioni tra task e fattori di processing psicolinguistico. Rispetto al primo elemento, la priorità è l'identificazione delle caratteristiche di un task in grado di avere un impatto positivo sulla qualità e complessità del processing linguistico. Il secondo elemento ha a che vedere con l'impatto della selezione dei task sull'uso linguistico, qui l'attenzione è posta sulle modalità con cui la performance porta a cambiamenti nella proficiency. Infine, il terzo elemento si rifà alla relazione tra task e fattori di processing, includendo pertanto temi come la natura del processo di comprensione, i modi in cui l'interazione porta l'apprendente a concentrarsi sulle relazioni forma-significato o come le caratteristiche dei task possano influenzare l'elaborazione linguistica.

L'interesse pedagogico per i task ha pertanto stimolato ricerche di vario tipo, raggruppabili almeno secondo quattro approcci principali: (a) l'approccio interazionale, influenzato principalmente dal lavoro di Long (1985; 1989); (b) l'approccio socioculturale, rappresentato dal lavoro di Lantolf (2000) e Swain (Swain 1998; Swain & Lapkin, 2001); (c) l'approccio focalizzato all'elicitazione di determinate strutture (*structure focused approach*), in cui i task sono impiegati per stimolare la produzione di forme linguistiche specifiche (Loschky & Bley-Vroman, 1993; Van Patten, 1996); e infine (d) l'approccio cognitivo di Skehan (Skehan, 1998; 2001; 2003; Skehan & Foster, 1999; 2001) e Robinson (2001a; 2001b; 2003; 2005; 2007).

A questo proposito è bene evidenziare come le ricerche basate sui task per l'insegnamento condotte fino ad oggi risentono di un limite: raramente inquadrano i task entro un contesto pedagogico più ampio, più frequentemente li studiano a prescindere dal contesto classe. Se questo ha permesso di approfondire le conoscenze rispetto allo sviluppo delle L2, in futuro sarà necessario porre maggior attenzione ai contesti pedagogici d'uso (Bygate *et al.*, 1998; Samuda, 2001).

La ricerca si è interessata ai task non solo come costruito, ma anche come strumento di indagine. In questo secondo caso si affronta il problema di come il task possa essere impiegato per osservare in modo ecologicamente più valido il processo di apprendimento. L'elicitazione dei dati e la loro analisi sono il punto di interesse principale e la selezione dei task è strettamente legata alle domande di ricerca. In questo caso vengono definiti task attività di vario tipo, in molti casi riferendosi a varie tecniche di elicitazione semi-spontanea (cfr. Gass & Mackey, 2007).

Anche nell'ambito del testing, in modo simile alla ricerca, l'interesse è rivolto verso task intesi come attività ben definite e limitate che rappresentino interazioni comunicative con una corrispondenza con i bisogni linguistici dell'apprendente, che possano essere valutate in modo valido e coerente e quindi raggiungere un certo grado di standardizzazione. Come sottolineano anche Bachman & Palmer (1996), se le caratteristiche di un task influenzano il risultato del test, allora diventa essenziale capire gli effetti di tale influenza. L'interesse è dunque rivolto in primo luogo alle diverse caratteristiche dei task che possono avere impatti significativi sulla performance, ed essere pertanto elementi fondamentali per l'interpretazione del test e la sua standardizzazione. La validazione d'altronde è un aspetto chiave nel testing (cfr. 2.3). I risultati ottenuti da un test possono infatti essere influenzati da una serie di elementi: valutatore e criteri di valutazione, condizioni di somministrazione e caratteristiche dei task. Le ricerche condotte in questo ambito hanno pertanto studiato fattori quali valutatori e scale di valutazione (per es. Lumley & McNamara, 1995; North, 1996), o gli effetti delle relazioni di potere tra

partecipanti (Van Lier, 1989) o le difficoltà di un task (Skehan, 2001). Quest'ultimo aspetto è stato affrontato secondo due punti di vista: utilizzando scale di valutazione (e.g. Wigglesworth, 1997; 2001; Chalhoub-Deville, 2001) o misure oggettive di sviluppo CAF (e.g. Skehan, 1998; Robinson, 2001). Ricerche di questo tipo possono offrire indicazioni estremamente utili a chi utilizza i task ai fini della valutazione.

Interessante notare come l'attenzione per la competenza comunicativa negli anni '70, il *proficiency movement* degli anni '80 e i più recenti stimoli nei confronti del *performance testing* abbiano portato chi si occupa di testing a condividere sempre più elementi con la didattica: il task rappresenta un esempio di questa maggior connessione tra valutazione e insegnamento. Dal punto di vista della metodologia di ricerca inoltre, le domande poste dall'impiego dei task richiedono una maggior interazione tra testing e ricerche acquisizionali (cfr. § 2.4). In questo studio si è cercato di coniugare le due prospettive, il task è qui usato come strumento di elicitazione, con l'obiettivo però di offrire elementi utili sia rispetto al tema della selezione di task per l'insegnamento o il testing, sia rispetto a quello della selezione di criteri affidabili per la misurazione della performance.

## **CAPITOLO 3**

### **METODOLOGIA DI RICERCA**

Questo capitolo presenta prima i soggetti dello studio, le modalità di raccolta dati e le caratteristiche del corpus, poi illustra le unità impiegate per l'analisi della produzione linguistica.

I dati considerati provengono da un ampio corpus realizzato con l'obiettivo di indagare la variabilità situazionale dell'interlingua di apprendenti avanzati di italiano L2. La raccolta è stata realizzata nell'ambito di due progetti di ricerca: COFIN 2003, co-finanziato dal MIUR e dall'Università di Verona (coordinatore locale Camilla Bettoni), e COFIN 2006, co-finanziato dal MIUR e dall'Università di Modena e Reggio Emilia (coordinatore locale Gabriele Pallotti), con il coordinamento in entrambi i casi a livello nazionale di Giuliano Bernini. Ho partecipato attivamente alla realizzazione di tale corpus collaborando alla definizione del protocollo, alla raccolta dati e all'organizzazione del database.

### 3.1. I soggetti

I soggetti presi in considerazione per questo studio sono sei. Si tratta di quattro apprendenti di italiano L2 a livello avanzato e due studentesse italiane. Al momento della prima rilevazione nel 2005 tutte e sei sono iscritte a istituti professionali, e hanno un'età compresa tra i 15 e i 19 anni. Le apprendenti risiedono in piccoli comuni della provincia tra Modena e Bologna, sono di diversa nazionalità e hanno frequentato con buon rendimento la scuola italiana per un periodo da 4 a 6 anni. Nei primi mesi in Italia avevano partecipato nelle rispettive scuole di arrivo a corsi di italiano L2 che mi hanno visto coinvolta come docente. La conoscenza e il rapporto di familiarità con i soggetti ha favorito la loro partecipazione attiva alla ricerca nel corso dei tre anni, oltre che reso più spontanee le interazioni. La tabella 6 riassume i dati dei sei soggetti. I paragrafi che seguono riportano invece un profilo delle apprendenti e delle studentesse italiane ricostruito sulla base della conoscenza diretta dei soggetti e dei racconti emersi durante le interviste.

**Tabella 6. I Soggetti**

	Pandita	Catherine	Eden	Shirley	Elisa	Valentina
codice	PAN	CAT	EDE	SHI	ELI	VAL
paese d'origine	India	Ghana	Eritrea	Nigeria	Italia	Italia
L1	punjabi	twi	tigrigno	inglese	italiano	italiano
età al t1	17	19	19	15	15	15
anni in Italia al t1	4	6	6	6		
livello QCER al t1 <sup>1</sup>	B1	B1	B2	B2	C1	C1

Pandita, 17 anni al momento della prima raccolta dati, è arrivata dall'India nell'estate del 2001 insieme alla sorella, al fratello e alla madre per ricongiungersi con il padre, che viveva già in provincia di Modena da una quindicina d'anni. Al suo arrivo Pandita viene iscritta alla scuola media del

<sup>1</sup> La valutazione del livello di competenza al t1 per i 6 soggetti è stata effettuata per altri scopi scientifici dall'equipe di ricerca COFIN utilizzando i descrittori globali del QCER.

comune di residenza e le viene richiesto di ripetere la classe terza, appena conclusa in India. Frequenta per tutto il primo anno scolastico corsi specifici di italiano L2 e gli insegnanti realizzano diversi progetti per facilitare la sua partecipazione alle lezioni delle diverse discipline. In generale Pandita ricorda positivamente l'impatto con gli insegnanti e con la scuola, nonostante le difficoltà oggettive dell'apprendimento dell'italiano e dell'incontro con la nuova realtà. Nella fase iniziale segnala come poco significativi i rapporti con i compagni, giustificando tale situazione sulla base delle difficoltà di comunicazione. Dopo aver superato con successo l'esame di licenza media, Pandita su consiglio degli insegnanti si iscrive ad un Istituto Tecnico Industriale a Modena. Frequenta la scuola con successo, anche se non nasconde le difficoltà affrontate nei primi anni: la lingua per lo studio continua a richiedere impegno e i rapporti con i compagni non sono molto intimi. Al momento della prima registrazione Pandita frequenta il terzo anno, in questa fase si sente più sicura rispetto alla lingua italiana e più autonoma nello studio, racconta inoltre di aver allacciato diversi rapporti con i compagni di scuola e valuta la sua esperienza scolastica come soddisfacente. Il recente viaggio in India sembra aver dato motivazioni nuove rispetto alla permanenza in Italia e Pandita si vede orientata verso una continuazione degli studi. Gli insegnanti la descrivono come una studentessa diligente e precisa. A t4 frequenta con soddisfazione e successo il primo anno di università.

Catherine, 19 anni al momento della prima raccolta dati, è arrivata dal Ghana all'inizio del 1999 per ricongiungersi con il resto della famiglia che viveva già da una quindicina d'anni in Italia. Catherine è la figlia più piccola di cinque fratelli. Al suo arrivo viene iscritta alla scuola media, per età avrebbe dovuto frequentare la terza, ma viene inserita in prima. I primi mesi di scuola frequenta un laboratorio di italiano L2. Catherine ricorda il primo anno come estremamente difficile, l'impatto con la nuova realtà e in particolare la lingua italiana non sono stati del tutto positivi. Nel giro di un paio d'anni riesce però a superare le difficoltà iniziali e comincia ad interagire più efficacemente con insegnanti e compagni. Al momento della prima registrazione, frequenta la



terza superiore di un Istituto Professionale ad indirizzo aziendale della provincia di Bologna. In questa fase giudica positivamente la sua partecipazione alla vita scolastica, segue nelle diverse discipline percorsi facilitati e dedica molto tempo allo studio. Catherine è piuttosto soddisfatta della sua vita in Italia, dove vede il suo futuro. A t4 frequenta con discreto successo il primo anno di università.

Eden, 19 anni al momento della prima raccolta dati, è arrivata dall'Eritrea nel 1999 insieme alla sorella per raggiungere la madre, che viveva da una decina d'anni in Italia. Fin da subito l'arrivo in Italia viene vissuto come rottura con il passato: la madre richiede l'uso della lingua italiana in casa e spinge entrambe le figlie ad inserirsi il più efficacemente possibile a scuola e nel territorio. Al suo arrivo Eden viene iscritta in prima media, anche se per età avrebbe dovuto frequentare la terza. Ricorda i primi anni di scuola come difficili, l'apprendimento della lingua italiana è stato inizialmente lento e impegnativo e i rapporti con i compagni non sono stati sempre facili, in alcuni casi addirittura conflittuali. Per Eden ha sempre costituito una grossa difficoltà la differenza di età, e di conseguenza di interessi e obiettivi, che la distingue dai compagni. Questo elemento ha caratterizzato l'intero percorso scolastico. A t1 Eden frequenta la classe terza di un Istituto Professionale ad indirizzo aziendale della provincia di Bologna. Gli insegnanti la descrivono come una studentessa seria e capace. Tra t2 e t3 Eden accetta una proposta di lavoro e continua la scuola seguendo i corsi serali. Il cambiamento è vissuto in modo estremamente positivo, in particolare rispetto al rapporto con compagni e colleghi. A t4 Eden ha conseguito il diploma e continua con soddisfazione la sua attività lavorativa.

Shirley ha 15 anni alla prima raccolta dati. È arrivata dalla Nigeria all'inizio del 1999 insieme alla sorella, per riunirsi con la famiglia, genitori e due fratelli, che risiedevano in Italia da più di 12 anni. Al suo arrivo viene inserita in seconda media. Durante i primi due anni di scuola frequenta corsi di italiano L2 e gli insegnanti di classe organizzano diversi progetti educativi per favorire il suo inserimento. Shirley ricorda positivamente il primo periodo a

scuola, in particolare il rapporto con gli insegnanti. Terminata la scuola media si iscrive ad un Istituto Professionale ad indirizzo turistico di Modena: è contenta della scelta fatta e partecipa con entusiasmo a stage o proposte formative extracurricolari offerte dalla scuola. A t1 Shirley frequenta la classe terza con buon rendimento, ha ottimi rapporti con i compagni e una buona rete di relazioni sociali. Frequenta assiduamente la chiesa e le attività della comunità ghanese di Modena. A t4 Shirley frequenta con successo il primo anno di università.

Elisa al momento della prima registrazione ha 15 anni ed è iscritta al secondo anno di un Istituto Professionale ad indirizzo aziendale. Vive in provincia di Bologna con i genitori ed un cagnolino. La scelta della scuola superiore si è basata più sulla posizione geografica della scuola che sulle reali propensioni di Elisa, che avrebbe preferito seguire un percorso di studi più legato alle arti. Nonostante questo il rendimento scolastico è discreto. Elisa ha gusti e preferenze molto precisi, sia nella scelta degli amici che delle attività da svolgere. Nel tempo libero si dedica alla musica, è inserita in una larga rete sociale di relazioni e si ritiene piuttosto soddisfatta di quello che fa. Gli insegnanti la descrivono come una studentessa riservata, ma capace.

Valentina al momento della prima registrazione ha 15 anni e frequenta il secondo anno di un Istituto Professionale ad indirizzo aziendale. Vive in provincia di Bologna con i genitori e un fratello più piccolo. L'esperienza scolastica di Valentina, nonostante risultati discreti, non è del tutto positiva. Gli insegnanti la descrivono come una studentessa capace, ma svogliata e demotivata. Vedremo che anche nella partecipazione a questo studio, la seconda rilevazione della sua produzione linguistica non è esente da problemi.

### 3.2. *Le rilevazioni*

I dati impiegati in questo studio sono stati raccolti utilizzando un protocollo di ricerca sperimentale per l'elicitazione di parlato semi-spontaneo elaborato dal

gruppo di ricerca coinvolto nei due progetti COFIN 2003 e COFIN 2006 (Pallotti, Ferrari & Nuzzo, in preparazione).

Tale protocollo è costituito di due momenti, indicati come giornata 1 e giornata 2. I task impiegati nella giornata 1 si rifanno a procedure tradizionalmente usate nella ricerca sull'acquisizione nella seconda lingua e coinvolgono l'intervistatore e un apprendente; i task utilizzati nella giornata 2 sono task comunicativi inseriti all'interno di un gioco a coppie a cui partecipano l'intervistatore, due apprendenti e diversi esperti adolescenti o adulti. La giornata 1 è costituita di quattro attività. La prima è un'intervista informale tra intervistatore e apprendente che segue una traccia tematica, anche se con una certa flessibilità. La seconda attività è il task *il racconto di un film*. La terza attività è il *map task* con adulti. In questo caso si è selezionato quello utilizzato nel progetto API (Archivio del Parlato Italiano, [www.cirass.unina.it](http://www.cirass.unina.it)). L'ultima attività è un task di *racconto di storia per immagini*. La giornata 2 è organizzata in tre task comunicativi: il *map task* con compagno, e due task in cui le studentesse per svolgere il compito assegnato – ad esempio, selezionare il miglior cellulare da un catalogo, organizzare la gita di classe, scegliere un libro o un dvd da regalare – devono raccogliere informazioni consultando telefonicamente esperti e negozi e concordare poi insieme una possibile soluzione.

Le produzioni delle apprendenti sono state registrate in quattro rilevazioni: t1 nel 2005, t2 nel 2006, t3 nel 2007, e infine t4 nel 2008. In t1 e t4 sono stati utilizzati gli stessi task, in t2 e t3 sono stati utilizzati task confrontabili. Le studentesse italiane sono state osservate invece in due rilevazioni: t1 nel 2005 e t2 nel 2007. Anche se nel loro caso sarebbe probabilmente bastata una sola rilevazione, si è comunque preferito effettuare un'ulteriore verifica.

### 3.3. *I task*

Il corpus di questo studio è costituito da quattro task: due task monologici, il racconto di un film e il racconto di una storia ad immagini, e due task interattivi, l'intervista e l'apertura delle telefonate di servizio, realizzati in quattro registrazioni effettuate nell'arco di tre anni.

Nel task racconto di un film, alle studentesse viene richiesto di guardare un video, della durata di 10 minuti circa e raccontarne poi la storia all'intervistatore, che non ne conosce il contenuto. Le studentesse guardano l'estratto una sola volta. Il materiale non è loro familiare e non è stato imposto un limite di tempo. Per questo task sono stati utilizzati il breve estratto *Alone and hungry* tratto dal film *Modern Times* di Charlie Chaplin in t1, due episodi della serie animata *Pink Panther* intitolati *Slink Pink* e *The Island* in t2 e t3 e nuovamente l'estratto dal video di *Modern Times* in t4.

Nel task racconto di una storia ad immagini, alle studentesse viene richiesto di narrare la storia all'intervistatore. Il materiale non è loro familiare e anche in questo caso non è posto un limite di tempo. In t1 è stato utilizzato il racconto ad immagini *Frog, where are you?* di Mayer, in t2 il racconto *Ho trovato un pettirosso* di Blanch, in t3 *One frog too many* di Mayer, e in t4 nuovamente il racconto *Frog, where are you?*.

Il task intervista consiste in una conversazione informale. Sono stati considerati due estratti per ciascun soggetto. Nel primo le studentesse parlano di se stesse e della propria famiglia, nel secondo raccontano del loro primo giorno di scuola in Italia o all'istituto professionale, o ancora dell'esperienza all'università o al lavoro. Il contenuto della conversazione viene mantenuto il più possibile costante tra t1 e t4. In ogni rilevazione però è stata coinvolta una persona diversa, non conosciuta dalle ragazze. Questa è sempre un'insegnante di italiano L2, di età compresa tra i 34 e i 38 anni, con una certa esperienza in attività di questo tipo. In questo modo le studentesse non hanno percepito il task come ripetitivo, ed è stato possibile mantenere sia un buon grado di spontaneità sia un buon livello di comparabilità tra rilevazioni.

L'apertura delle telefonate di servizio è un task incluso in diversi task comunicativi più ampi del corpus che prevedono, tra l'altro, una serie di telefonate per raccogliere informazioni ed accordarsi sulla scelta del miglior cellulare, libro, cd musicale o film e sull'organizzazione di una gita scolastica o un viaggio per sé o altri. Questi task sono dunque a soluzione aperta: per svolgere il compito assegnato le apprendenti devono raccogliere informazioni consultando telefonicamente varie persone sconosciute, tra cui coetanei, adulti e personale di negozi specializzati e concordare infine la soluzione ottimale. In questo studio sono state considerate da cinque a sette telefonate per soggetto, analizzandone esclusivamente l'apertura perché rappresenta un (sotto)evento comunicativo relativamente uniforme e nel quale le partecipanti producono turni di una certa lunghezza.

### 3.4. *I dati*

Le produzioni realizzate dai sei soggetti nei quattro task eseguiti per quattro volte dalle apprendenti e due dalle italiane, sono state videoregistrate, trascritte e segmentate in AS-unit, l'unità di analisi del parlato ritenuta più adatta agli scopi dello studio (cfr. § 3.4.1). Il corpus della tesi è pertanto costituito da 35.710 parole, equivalenti a 6.664 AS-unit (cfr. tabella 7).

Per la trascrizione si è fatto riferimento al sistema jeffersoniano (Atkinson & Heritage, 1984) e a CHAT ([www.childes.psy.cmu.edu](http://www.childes.psy.cmu.edu)), con alcuni adattamenti necessari per il trattamento informatico dei dati: £parola£ indica una parola pronunciata ridendo; .hhh e hhh indicano rispettivamente inspirazione ed espirazione; paro:::la segnala un prolungamento di suono; f\*parola\*f e s\*parola\*s segnalano rispettivamente parlato veloce e parlato lento; il punto fermo indica intonazione finale discendente, la virgola intonazione finale sospesa ed il punto interrogativo intonazione finale ascendente; # segnala le micropause, #2\_0 indica la durata della pausa, in questo caso 2 secondi.

Per la segmentazione in unità di produzione, le AS-unit sono contenute tra barre verticali doppie (**|| xxxxx ||**), mentre le clausole o unità inferiori alla clausola sono racchiuse tra barre verticali semplici (**| xxxxx |**). Infine, false partenze o ripetizioni vengono racchiuse tra parentesi graffe.

**Tabella 7. Il Corpus**

	<i>apprendenti</i>				<i>italiane</i>		totale
	t1	t2	t3	t4	t1	t2	
<b>parole</b>							
film	1982	1480	1562	2955	844	552	9375
storia	2209	1869	1846	2595	1057	661	10237
intervista	3014	3238	3946	1290	1254	662	13404
telefonata	361	461	614	691	332	235	2694
totale	7566	7048	7968	7531	3487	2110	35710
<b>AS-unit</b>							
film	362	253	248	515	124	75	1577
storia	419	312	281	420	167	91	1690
intervista	650	630	735	225	248	259	2747
telefonata	95	111	159	169	73	43	650
totale	1526	1306	1423	1329	612	468	6664

### 3.5. *L'unità e le misure dell'analisi*

Questa sezione illustra gli strumenti di misurazione utilizzati per l'analisi del mio corpus, prima quelli del CAF (§ 3.5.1), poi quelli del QCER (§ 3.5.2). Poichè qualsiasi strumento di misurazione, oltre che valido, deve essere anche affidabile, la presentazione dettagliata delle unità di analisi è accompagnata quando possibile da una verifica sperimentale della loro affidabilità.

### 3.5.1. Le misure CAF

Per misurare CAF nelle produzioni sono state proposte diverse misure quantitative come abbiamo visto in § 2.2.2.2. In questa sezione illustro quelle scelte per questo studio e ne giustifico l'utilità. Dapprima presento l'unità di produzione, specificando gli adattamenti resi necessari per la sua applicazione a dati italiani, poi presento le misure quantitative CAF.

Foster *et al.* (2000) definiscono l'AS-unit, ovvero l'*Analysis of Speech Unit* come “a single speaker's utterance consisting of an *independent clause*, or *sub-clausal unit*, together with any *subordinate clause(s)* associated with either.” In (3) ne diamo un esempio:

- (3) SH1: || | e::: la pantera? #0\_4 {e:::} ha sempre fatto di tutto. | in modo che:::  
{il ca} il cacciatore s::i::: incavoli con e::: il cane. || #0\_2

Questi autori distinguono poi le clausole che costituiscono le AS-unit in:

- clausole indipendenti
- unità indipendenti inferiori alla clausola
- clausole subordinate.

Una clausola indipendente è definita come “a clause including a finite verb”:

- (4) ELI: || | e::: {va:::} ruba {un pezzo di legno} # insomma del fuoco:: dalla casa  
del cacciatore # || 0\_6.hhh

Un'unità indipendente inferiore alla clausola “will consist of either one or more phrases which can be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse or situation [...] or a minor utterance, which will be defined as one of the class of 'Irregular sentences' or non sentences identified by Quirk *et al.*”:

- (5) SH2: sì  
SH1: || | allora Danielle Steel | ||  
SH2: sì  
SH1: || | il ranch | ||  
SH2: vediamo

Una clausola subordinata “will consist minimally of a finite or non finite verb element plus at least one other clause element (subject, object, complement or

adverbial)”:

- (6) PAN: || mi puoi prendere. #0\_3 | perché ho mangiato tutto. ||

Le false partenze, le ripetizioni e le autocorrezioni vengono escluse dalla segmentazione in AS-unit e pertanto figurano tra parentesi nella trascrizione. Con falsa partenza Foster *et al.* si riferiscono sia ai casi in cui il piano discorsivo viene abbandonato sia ai casi di riformulazione. Se una AS-unit o una clausola completa è prodotta prima che il piano venga abbandonato, essa viene invece inclusa nel conteggio:

- (7) ELI: || lui s- muove un: e:: || #1\_6  
ELI: || non lo so. ||  
ELI: || comunque lui lavora in un cantiere tipo ||

Nel caso di ripetizioni si distingue tra ripetizione come disfluenza (8), quindi esclusa dall’analisi, e ripetizione come strategia retorica (9), quindi inclusa nel conteggio:

- (8) PAN: || e- adesso quando {il >polisiotto} anche il polisiotto< è alzato? ||  
#0\_2  
PAN: || ha dato una #0\_4 botta nella testa? ||  
INV: £hh£  
PAN: || .h poi {sono cs scappati? #0\_3 e: #0\_6} s:ono scappati ||  
PAN: || e sono .h seduti una vicina casa? ||
- (9) PAN: || poi Charlot- entra in <caffettiera> ||  
PAN: || e mangia tutto tutto ||  
PAN: || non paga niente? ||

Si parla di autocorrezione quando un parlante identifica un errore durante o subito dopo la sua produzione e intende correggerlo riformulando. Quando ciò accade si conteggia l’ultima formulazione:

- (10) VAL: || ed era {dafra-} davanti {a un::} a #0\_2 un fornaio così ||

Si escludono dal conteggio *tch*, *mh* o segnali discorsivi come *cioè*:

- (11) VAL: || {cioè} si capisce || #0\_5  
VAL: || {cioè} non lo dice {cc-} ||  
INV: da cosa si capisce h  
VAL: || {boh} non lo so ||

Nell’interazione sono frequenti anche fenomeni di *scaffolding* e interruzione. Nei casi in cui una subordinata sia collegata alla produzione precedente dell’apprendente, questa verrà considerata come clausola della



stessa AS-unit, anche se conclusa dopo un turno dell'interlocutore:

- (12) PAN: || [scappa dietro: #0\_5 alla ragazza #0\_5 | ||  
PAN: || e: la ragazza #0\_3 cade sopra: h #0\_7 il vuoto .h | che l- era andato |  
a lavorare {in: °n na- # } in nave per la°  
INV: .h ne nel cantiere  
PAN: cantiere ||

Nonostante l'AS-unit come unità sia stata operativizzata con notevole precisione, essa è stata proposta per l'analisi della lingua inglese, pertanto la sua applicazione a dati orali in italiano non è scontata. Applicando l'unità al mio corpus sono emerse due tipi di questioni, la prima rispetto ai confini esterni all'unità: 'una o due AS-unit?'; la seconda rispetto ai confini interni: 'una o due clausole?'. Qui di seguito, quindi, prima illustro le precisazioni che ho apportato alla definizione operativa di AS-unit, e poi descrivo la procedura impiegata per verificarne l'affidabilità nell'applicazione.

L'identificazione del confine dell'unità non è sempre scontata. Anche Foster *et al.* (2000) riscontrano alcuni problemi nel riconoscimento della coordinazione tra sintagmi verbali:

There is the problem of coordinated verb phrases. We argue that in many cases, the second VP, even if the subject is omitted, constitutes a fresh start for the speaker. To address this problem we have identified conditions for determining where the unit boundary lies. In cases where coordination of verbs occurs, the coordinated phrases will normally be considered to belong to the same as-unit, unless the first phrase is marked by falling or rising intonation or is followed by a pause of at least 0.5 seconds.

Seguendo queste indicazioni (13) e (14) sarebbero analizzati così:

- (13) ELI: || e:: solo che esplode tutto | e viene un buco nell'isola? || #0\_6  
ELI: || e {sprofonda::} fhhh\$ sprofonda nel mare | ||  
(14) SHI: || {mch} arriva la: polizia::? | ||  
SHI: || e::: .h # l'uomo # dice. | ||

Come applicare però la stessa definizione ai seguenti estratti?

- (15) PAN: non ha rubato {nan} #0\_4 un vuoto? # il pane. {h } ha rubato una donna?  
(16) SHI: e: # l'ha sempre bastonato::: gli ha:: sparato::: gli ha fatto di tutto. però::? {alla::} alla # fine::: {il} il cane viene # f\*sbattuto fuori\*f .perché:: {il} il cacciatore non ne poteva più. e::: l'ha cacciato fuori? e: lui ha::: cominciato:: a ri ribellarsi nei confronti dei::: del # cacciatore.

Determinare l'esistenza di un legame sintattico affidandosi a criteri di tipo intonativo o alle pause potrebbe limitare l'affidabilità della misura. Infatti il ruolo delle pause e dell'intonazione non è sempre univoco: perciò non è sempre possibile decidere ad esempio se una pausa sia conseguenza della ricerca di una parola o della pianificazione di un nuovo enunciato.

Dal punto di vista sintattico, poi, nonostante sia chiaro che la coordinazione è una relazione di non dipendenza, rispetto alla dipendenza della subordinazione, come viene sottolineato nella letteratura sulla tipologia dei legami inter-clausali (cfr. Talmy, 1978; Longacre, 1985; Lehmann, 1988), coordinazione e subordinazione possono essere descritte come i due poli di un continuum. A un estremo non troviamo nessuna relazione gerarchica tra le clausole che compongono una frase complessa, dall'altro troviamo un chiaro rapporto di dipendenza, dove la clausola subordinata è un costituente ben definito della frase. La posizione di una clausola subordinata rispetto alla reggente è legata a vincoli di tipo grammaticale. Questo non avviene nel caso della coordinazione: se invertiamo la posizione di due clausole coordinate possiamo forse confondere la chiarezza semantica dell'enunciato, ma la sintassi rimane indifferente all'ordine delle clausole. Un'unità di produzione quantitativa per essere affidabile dovrà invece avere linee di demarcazione precise. Qualunque scelta presenta dunque un certo grado di arbitrarietà e nel mio caso si è cercato di favorire al massimo un'applicazione affidabile della misura.

Avendo in mente le caratteristiche del mio corpus e gli obiettivi della mia ricerca, tra cui l'esplorazione della complessità sintattica intesa come capacità di produrre frasi complesse, ho preferito scostarmi dalle indicazioni offerte da Foster *et al.* (2000), e fissare la linea di demarcazione tra coordinazione e subordinazione piuttosto presto, considerando le clausole coordinate come AS-unit indipendenti. In questo modo l'AS-unit risulta più affidabile nella sua applicazione e maggiormente legata a principi sintattici.

Un altro punto critico rispetto alla definizione dei confini esterni all'AS-unit riguarda l'interpretazione di alcuni tipi di clausola introdotti da *perché* o

*che*, a volte chiaramente interpretabili come subordinate, a volte come nuovo inizio:

- (17) VAL: || | prendete me | perché son stata io. | || #0\_4
- (18) PAN: || | [era f-]cane era fuori. | ||  
 PAN: || | perché::: dopo che l'hanno cacciato fuori. |  
 INV: eh.  
 PAN: | guardava dalla finestra. | || # 0\_4

In (17) è piuttosto evidente il legame di dipendenza sintattica oltre che semantica del connettore causale *perché*, supportato dall'intonazione, poichè la subordinata causale è inclusa nello stesso profilo intonativo della reggente. In (18) invece *perché* ha più le funzioni di un segnale discorsivo, e la nuova AS-unit viene collegata semanticamente a ciò che è stato detto, ma non sintatticamente. In questo caso le due clausole sono distribuite su due profili intonativi distinti e *perché* è caratterizzato da un allungamento della vocale finale. L'intonazione in questi casi aiuta la scelta del ricercatore. Analoghi i due esempi con *che*:

- (19) SHI: || | trova: {il: # e:il: .h #0\_5 } l'uomo | che sta facendo # il pane. | che dal furgone va dentro {a a} al negozio? | ||
- (20) SHI: || poi: e:: mentre .h # stava ancora pensando | sono tornati poi nella:: .hhh realtà | dove: # arriva # il poliziotto. | ||  
 SHI: || | che::: cominciano ancora a scappare. | ||

Anche in questo caso *che* può rappresentare il connettore di una relativa (19) o essere sintatticamente inattivo (20) e avere dunque un effetto di subordinazione solo semantica con quanto precedentemente detto. Nel caso del *che* relativo, la subordinata è sotto lo stesso profilo intonativo della reggente (o con profilo intonativo parallelo), mentre nel caso del *che* sintatticamente inattivo, le due clausole cadono sotto un diverso contorno intonativo e il *che* è caratterizzato da un allungamento vocalico (cfr. Cinque, 2001; Scarano, 1996; 2003). In entrambi i casi dunque il ricercatore può basarsi sull'esistenza di relazioni sintattiche ed essere poi guidato a scegliere da elementi intonativi.

Il secondo tipo di problema che emerge nell'applicazione della AS-unit così come definita da Foster *et al.* (2000) riguarda la determinazione dei confini interni alla AS-unit, con particolare riferimento alla determinazione

delle caratteristiche che una clausola deve avere per essere considerata subordinata.

Siccome per Foster *et al.* (2000), come abbiamo visto, una clausola subordinata “will consist minimally of a finite or non finite verb element plus at least one other clause element”, clausole avverbiali o relative non pongono particolari difficoltà:

- (21) VAL: || ha fatto di tutto | che così il pa#drone lo mandasse via. |||
- (22) VAL: || | tch h e:m dopo che::: l'hanno licenziato | in pratica: sta passeggiando lungo la strada h | quando vede: una ragazza #0\_2 | **che**: {sta ruba cioè} sta rubando un filoncino di pane |||

Ma come interpretare “plus at least one other element” nei casi che seguono?

- (23) PAN: || | è cominciato a mangiare: |||
- (24) PAN: || | mh # e::: quindi comincia: a boiare# ancora? || #0\_6
- (25) VAL: || | la voleva portare con sé in prigione |||
- (26) VAL: || | vallo a cercare || #0\_2

Seguendo la definizione di Foster *et al.* gli estratti in (23) e (26) dovrebbero essere analizzati come AS-unit di una clausola, quelli in (24) e (25) come AS-unit di due clausole. Questi esempi richiedono una definizione di clausola subordinata più precisa. Come devono essere considerati i verbi che dimostrano un alto grado di integrazione, come i modali e i fasali seguiti dall’infinito? La letteratura sull’italiano L1 e L2 mostra come non ci sia un accordo generale nelle modalità di analisi di queste strutture. Seguendo Giacalone Ramat (1999a; 1999b), gli esempi precedenti sarebbero clausole indipendenti seguite da una subordinata implicita. Seguendo Renzi *et al.* (1988-1995) e Chini (1998; 2003) sarebbero invece clausole indipendenti costituite da un verbo complesso.

Nella mia analisi pongo linee di demarcazione precise per distinguere i casi in cui il verbo reggente e l’infinito formano una specie di verbo composto da quelli in cui l’infinito costituisce un membro indipendente del sintagma verbale. Secondo Renzi & Salvi (2001: cap. 9) e Graffi (1994; 2001), un verbo reggente seguito da un’infinito costituisce un verbo complesso se sono

soddisfatte le seguenti condizioni: (a) i clitici, pur dipendendo dall'infinito si trovano accanto ai verbi reggenti; (b) il complemento oggetto del verbo all'infinito, può diventare il soggetto della costruzione con il *si* passivo; (c) l'infinito non può essere negato. Pertanto l'infinito non può essere considerato il predicato di una subordinata nei seguenti casi: (a) in costruzioni fattitive o causative, come *fare*, *lasciare* e alcuni verbi percettivi, *intendere*, *sentire*, *udire* e *vedere*:

(27) VAL: || e ci\_ h # aveva fatto scappare la ragazza prima h | ||

(28) SHI: || | pantera rosa. | #0\_4 che: ho:: sentito parlare. | non ho mai avuto ehm guardato ne[ssuna puntata del cartone | ||

e (b) in costruzioni a ristrutturazione, con verbi modali (*dovere*, *potere*, *volere* e *sapere*), verbi aspettuali (*cominciare*, *finire*) e verbi di moto (*andare*, *venire*):

(29) PAN: || | poi: #0\_5 la: ragazza ha cominciato d- a scappare | ||

(30) PAN: || | e::: la ragazza #0\_3 cade sopra: h #0\_7 il uomo .h | che l- era andato a lavorare {in: °n na- # } in nave per la°  
INV: .h ne nel cantiere  
PAN: cantiere | ||

(31) VAL: || | [e quindi vorremmo sapere | se lei ci può consigliare | ||

Anche le perifrasi verbali vengono considerate un'unica clausola:

(32) SHI: || | poi: e::: mentre .h # stava ancora pensando | sono tornati poi nella:: .hhh realtà | dove: # arriva # il poliziotto | ||

Un ulteriore aspetto critico della segmentazione in AS-unit riguarda alcune costruzioni marcate tipiche del parlato, che in particolare se dialogico è caratterizzato dall'effetto di sintassi fratta, con una struttura per periodi giustapposti. In primo luogo, il parlato si sviluppa attraverso microprogettazione tematica e sintattica: ne sono la manifestazione pause, esitazioni, piani abbandonati, frasi incompiute, nonché le repentine e frequenti fratture di continuità tematica non segnalate esplicitamente da connettivi, ma solo da pause e da demarcativi riempitivi e fatichi. In secondo luogo, le riprese e i mutamenti di topic spesso non vengono esplicitate, mentre le varie parti del discorso vengono cucite con segnali di articolazione, la catena anaforica si realizza attraverso riprese ad hoc, rimandi all'universo di discorso, realizzando

così la minima codifica morfosintattica. Infine frequente è l'uso di dislocazioni o frasi segmentate (Bazzanella, 1991; 1994; 2002).

Uno dei fenomeni più caratteristici della sintassi del parlato riguarda proprio la variazione nell'ordine delle parole, che in questa modalità segue principi di strutturazione diversi rispetto all'ordine canonico dell'italiano, in cui il soggetto è tema e dato e viene a sinistra del verbo, e l'oggetto è rema e nuovo e viene a destra. Particolari realizzazioni marcate dell'ordine dei costituenti sono connesse sia con la struttura spezzettata e disarticolata, sia con l'accentuazione empatica, tipiche del parlato. I costrutti marcati sono di tre tipi: dislocazione a sinistra, dislocazione a destra e frasi segmentate. Se i primi due non pongono particolari difficoltà nell'identificazione dei confini interni alle unità, le frasi segmentate in alcuni casi possono prestarsi ad interpretazioni discordanti. In questo studio si è pertanto stabilito che le frasi segmentate costituiscono una AS-unit formata da una sola clausola:

- (33) ELI: || | volevo chiederle informazione | || #  
 AG2: sì  
 ELI: || | e: per Londra | || #  
 AG2: sì  
 ELI: || | una classe di: ragazzi | ||  
 AG2: mh mh  
 ELI: || | e: qualcosa di conveniente {che::} # che c'è | ||
- (34) ELI: || | volevo chiederle un'informazione | ||  
 SH2: sì  
 ELI: || | lei non è che avrebbe per caso dei cd in offerta dei pink floyd | || #

Un'altra costruzione funzionalmente simile è rappresentata dalle presentative del tipo *esserci x che*:

- (35) PAN: || | c'era u:n bastone | #0\_3 che {c'è} >sotto< c'era una .h cono | #0\_5 che ha fermato nave? | ||

Questo tipo di costrutto è particolarmente utile e pertanto frequente nel parlato spontaneo con gittata di pianificazione ridotta, poiché il parlante trova vantaggio nell'enunciare un oggetto di discorso, lasciando al momento processuale successivo la predicazione su di esso. Allo stesso modo l'informazione distribuita su due frasi risulta di maggior agio ricettivo anche per l'ascoltatore. Berruto (1986) considera queste costruzioni pseudo-relative e

assegna loro un valore fondamentalmente esplicativo e di commento. Anche Cinque (2001: 502) le interpreta come frasi scisse piuttosto che relative. La frase introdotta da *che* non è infatti un modificatore di *x*, ma contiene la proposizione principale del costrutto, di cui il sintagma scisso è argomento necessario. Nonostante il valore semantico sia ben diverso da quello delle relative standard, esse condividono almeno in parte, certe caratteristiche di superficie (Bernini, 1991). Per questa ragione, in sede di analisi le costruzioni del tipo *esserci x che y* vengono interpretate come una AS-unit di due clausole. Interessanti però i due esempi (36) e (37), dove la frase presentativa è ellittica, con il *c'è* sottinteso, e la clausola indipendente viene pertanto codificata come unità indipendente inferiore alla clausola:

(36) SHI: || | allora mia sorella grande | che non c'è | ||

(37) SHI: || | il mio fratello | che: si chiama m[artin. | ||

Per verificare l'affidabilità della AS-unit è stato effettuato un controllo sperimentale su una parte del corpus, che ha coinvolto due valutatori. In una prima fase i valutatori hanno potuto familiarizzarsi con la definizione di AS-unit discutendo alcuni estratti dal corpus e esercitandosi nella segmentazione. In una seconda fase hanno, l'uno indipendentemente dall'altro, segmentato l'intervista, il racconto di un film e l'apertura delle telefonate realizzati da Pandita e da Elisa a t1 (1278 parole). Si è proceduto poi al confronto delle due codifiche (*inter-coder agreement*). Poichè in letteratura si considerano affidabili i giudizi quando il valore della correlazione è pari o superiore a 0,80 (Alderson *et al.*, 1995), l'applicazione della AS-unit così ridefinita ha dimostrato un ottimo grado di affidabilità, con un accordo tra i due valutatori nell'identificazione sia di clausole che di AS-unit pari al 98%.

La tabella 8 riassume le caratteristiche della AS-unit ridefinita per l'italiano e adattata al mio corpus.

**Tabella 8. L'AS-unit, Analysis of Speech Unit**

<i>unità</i>	<i>esempio</i>	<i>definizione</i>
clausola indipendente	ELI:    e:: {va::} ruba {un pezzo di legno} # insomma del fuoco:: dalla casa del cacciatore    VAL:    la voleva portare con sé in prigione    PAN:    è cominciato a mangiare:	clausola che include un verbo finito, anche complesso: costruzioni fattive e causative; con verbi modali, aspettuali, di moto; perifrasi verbali
unità indipendente inferiore alla clausola	SHI:    allora Danielle Steel    SH2: sì SHI:    il ranch    SH2: vediamo	clausola costituita da una o più parole rielaborabile come completa recuperando elementi sottintesi: profrasi, segni predicativi autonomi a nodo centrale non verbale, clausole nominali, di replica o troncate
clausola subordinata	PAN:    mi puoi prendere. #0_3   perché ho mangiato tutto.    VAL:    tch h e:m dopo che::: l'hanno licenziato   in pratica: sta passeggiando lungo la strada h   quando vede: una ragazza #0_2   che: {sta ruba cioè} sta rubando un filoncino di pane    PAN:    {h} c'è un:: # uomo   che dice   di prendere una: # legna {n} come {co;} co:no	clausole argomentali, avverbiali, relative. Si distinguono le costruzioni con verbi complessi dalle costruzioni infinitive vere e proprie
<i>fenomeno</i>	<i>esempio</i>	<i>analisi</i>
clausola coordinata	VAL:    vallo a cercare.       e lui va a cercarlo.	si conteggia come AS-unit indipendente
clausola introdotta da <i>che</i> o <i>perché</i>	VAL:    prendete me   perché son stata io.    PAN:    [era f-]cane era fuori.       perché::: dopo che l'hanno cacciato fuori.   INV: eh. PAN:   guardava dalla finestra.	si distinguono i connettori sintatticamente attivi dai casi in cui <i>che</i> o <i>perché</i> sono segnali discorsivi
frase segmentata	ELI:    volevo chiederle un'informazione    SH2: sì ELI:    lei non è che avrebbe per caso dei cd in offerta dei pink floyd    #	si considera come AS-unit formata da una clausola



<i>scaffolding</i> e interruzione	PAN:    e: la ragazza #0_3 cade sopra: h #0_7 il uomo .h   che l- era andato   a lavorare {in: °n na- # } in nave per la° INV: .h ne nel cantiere PAN: cantiere	se una clausola è collegata alla produzione precedente dell'apprendente, viene considerata parte della stessa AS-unit
ripetizione	PAN:    e- adesso quando {il >polisiotto} anche il polisiotto< è alzato?       PAN:    poi Charlot- entra in <caffettiera>         e mangia tutto tutto         non paga niente?	si distingue tra ripetizione come disfluenza, non inclusa nell'analisi, e ripetizione come strategia retorica, inclusa
autocorrezione	VAL:    ed era {dafra-} davanti {a un::} a #0_2 un fornaio così	si conteggia l'ultima versione
falsa partenza	ELI:    lui s- muove un: e::      #1_6    non lo so.        comunque lui lavora in un cantiere tipo	esclusa dall'analisi, a meno che una AS-unit o una clausola completa sia prodotta prima che il piano venga abbandonato
<i>tch, mhm</i> o segnali discorsivi come <i>cioè</i>	VAL:    {cioè} si capisce      VAL:    {cioè} non lo dice {cc-}      INV: da cosa si capisce h VAL:    {boh} non lo so	si escludono dal conteggio

Definita così l'AS-unit e stabilita la sua affidabilità, vediamo ora le misure CAF. Tra quelle proposte dalla letteratura (cfr. § 2.2.2.2) illustro le misure scelte per questo studio e ne giustifico l'utilità. Presenterò prima le misure relative alla complessità sintattica, poi quelle relative ad accuratezza, e infine illustrerò le misure di fluency.

Abbiamo visto in § 2.2.2.2 come le misure di complessità sintattica siano numerose e alternative. Le riassumo nella tabella 9. Esse rilevano due aspetti di questo processo di complessificazione: il grado, ossia la capacità di usare strutture più complesse sia a livello di clausola che di AS-unit; e la gamma, ossia la capacità di usare una varietà di strutture. Applicare tutte queste misure al corpus sarebbe oltre che oneroso in termini di tempo anche poco pratico, soprattutto in funzione dell'interpretazione dei risultati. Se in alcuni casi le informazioni fornite da queste misure certamente non sono uguali, in altri sono ridondanti. Per questo è stato necessario fare una cernita.

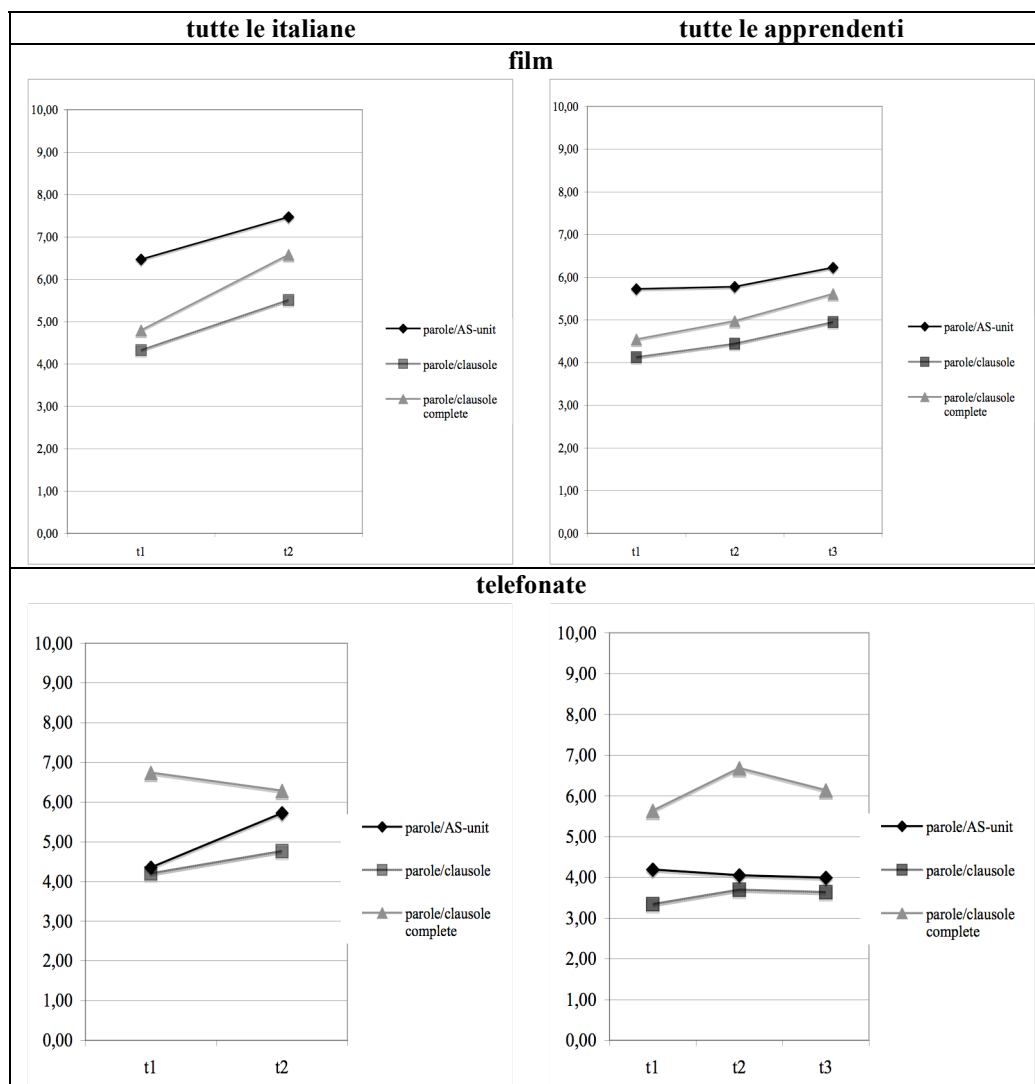
**Tabella 9. Le misure di complessità sintattica proposte dalla letteratura**

<i>misura globale di grado</i>	
lunghezza della AS-unit	numero medio di parole per AS-unit
<i>misura specifica di grado</i>	
lunghezza della clausola	numero medio di parole per clausola
	numero medio di parole per clausola completa (escludendo cioè le unità inferiori alla clausola)
grado di subordinazione	numero medio di clausole subordinate per AS-unit
	numero medio di clausole subordinate per clausola
	numero medio di clausole subordinate per clausola completa
<i>misura qualitativa di gamma</i>	
tipo di clausola	percentuale dei 3 tipi di clausole: indipendenti, unità indipendente inferiore alla clausola, dipendenti sul totale delle clausole
	percentuale dei tipi di subordinate: argomentali, avverbiali e relative sul totale delle subordinate

La misura globale di grado è rappresentata dalla lunghezza delle unità di produzione. Abbiamo visto in § 2.2.2.2 come non ci sia generale consenso in letteratura sulla funzione di questa misura, alcuni ricercatori la ritengono un indicatore della complessità sintattica, altri della fluenza. In questo studio questa misura viene considerata un indicatore globale della complessità sintattica. Le misure globali e specifiche di grado si distinguono tra di loro, in primo luogo per elemento al numeratore, ossia il numero di parole e di subordinate, e in secondo luogo per l'elemento al denominatore, ossia se si tratta di AS-unit, clausola o clausola completa. Possiamo infatti ipotizzare che una differenza nel numeratore generi misure complementari, una differenza nel denominatore generi probabilmente misure ridondanti (cfr. Norris & Ortega, in stampa). Per verificare questo sperimentalmente ho considerato il parlato prodotto dai sei soggetti nelle prime tre rilevazioni in due task del corpus: film e apertura delle telefonate. I risultati ottenuti dai due gruppi applicando le misure di grado sono stati poi confrontati. Il grafico 1 riporta le medie dei risultati ottenuti sulle misure di lunghezza per i due gruppi delle quattro

apprendenti e delle due studentesse italiane, mentre la tabella 10 riporta la variazione tracciata dalle misure.

**Grafico 1. CAF – Il confronto tra le misure di lunghezza – Italiane e Apprendenti (valori di gruppo)**



**Tabella 10. Le misure di lunghezza, variazione**

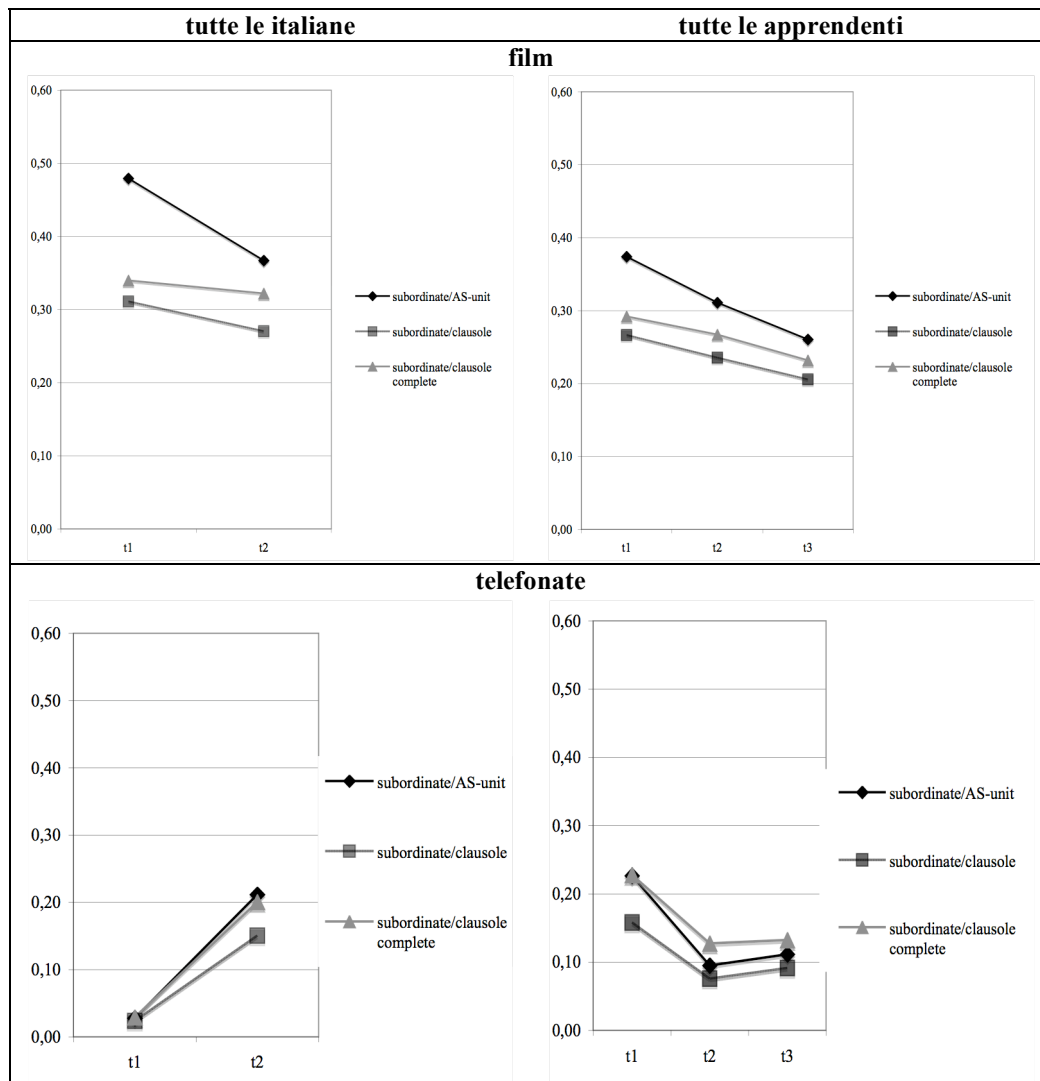
	<i>film</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
parole/AS-unit	8,70%	15,44%	-4,74%	31,55%
parole/clausole	19,87%	27,40%	8,73%	7,14%
parole/clausole complete	23,38%	30,39%	8,92%	7,12%

Confrontando gli andamenti di parole/AS-unit e parole/clausole o parole/clausole complete si osserva come l'andamento di parole/AS-unit sia diverso rispetto a quello delle altre due misure. Se si considerano ad esempio i valori ottenuti nel task apertura delle telefonate realizzato dalle apprendenti, è evidente come la misura parole/AS-unit abbia non solo un andamento diverso, ma addirittura inverso rispetto agli altri due: la misura parole/AS-unit rileva in questo caso una variazione in negativo del 4,74% tra t1 e t3, mentre le misure parole/clausole e parole/clausole complete variano in positivo dell'8,73% e dell'8,92% rispettivamente.

Questo supporta il dubbio già emerso in letteratura (cfr. § 2.2.2.2), che la misura basata sulla AS-unit e le misure basate sulla clausola rilevano cose diverse. La prima rileva la complessificazione generata sia a livello di clausola sia tramite l'uso di subordinate, le seconde rilevano invece solo il primo tipo di complessificazione. Confrontando invece le due misure parole/clausole e parole/clausole complete vediamo che hanno andamenti molto simili e quindi possiamo ritenerle ridondanti.

Il grafico 2 riporta le medie dei risultati ottenuti per le misure del grado di subordinazione dai due gruppi delle studentesse italiane e delle apprendenti, e la tabella 11 mostra la variazione tra misure. Qui constatiamo che le tre misure (subordinate/AS-unit vs subordinate/clausole e subordinate/clausole complete) hanno andamenti molto simili, con l'unica differenza che subordinate/AS-unit sembra un po' più sensibile al mutamento, dimostrando in tutti i contesti maggior variazione. Possiamo pertanto ritenere che le tre misure siano equivalenti.

**Grafico 2. CAF – Il confronto tra le misure di subordinazione – Italiane e Apprendenti (valori di gruppo)**

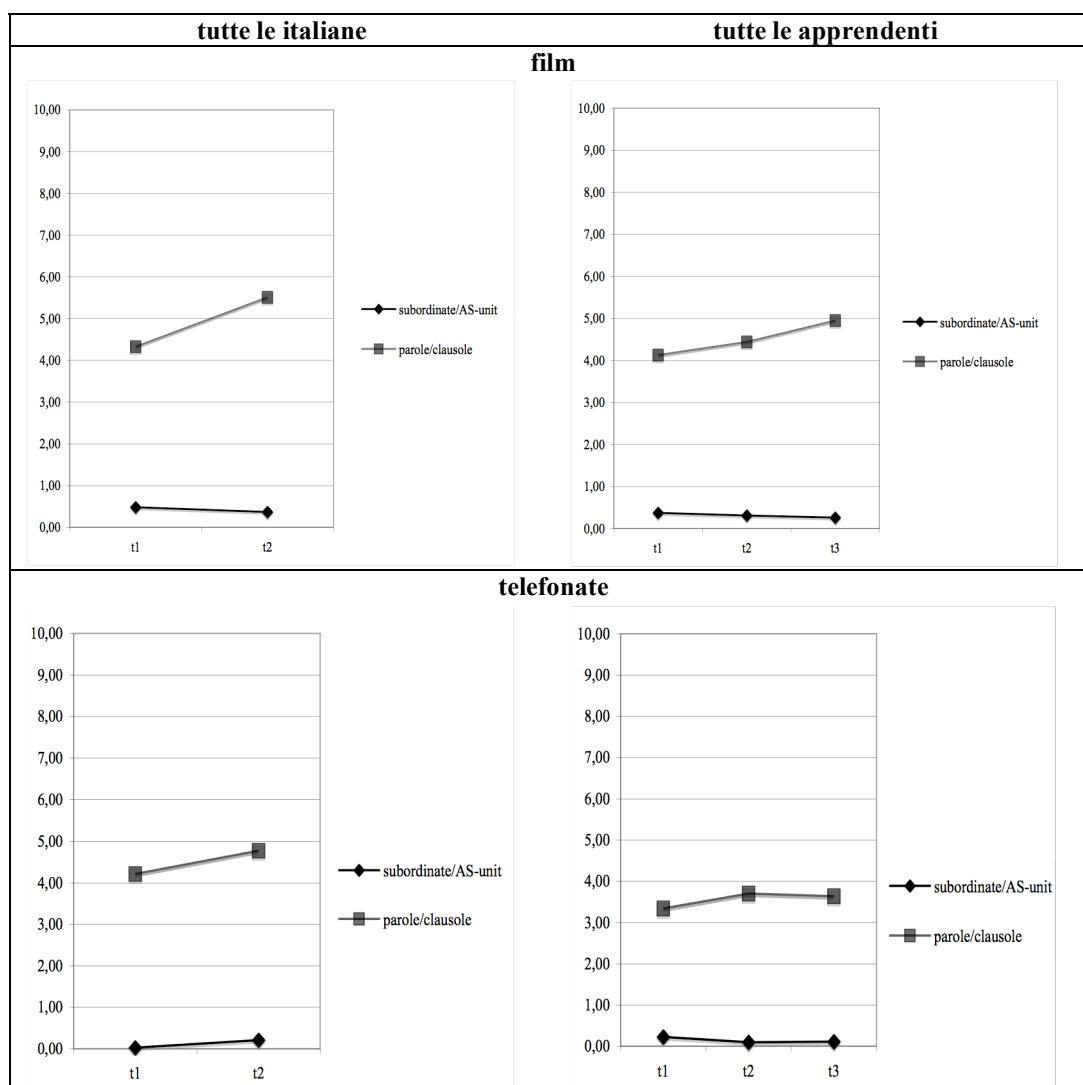


**Tabella 11. Le misure di subordinazione, variazione**

	<i>film</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
subordinate/AS-unit	-30,36%	-23,46%	-50,86%	11,34%
subordinate/clausole	-22,83%	-13,12%	-42,14%	8,92%
subordinate/clausole complete	-20,60%	-9,97%	-41,87%	12,04%

Sulla base di questi risultati ho deciso di utilizzare solo due misure per l'analisi: per la lunghezza la misura parole/clausole e per il grado di subordinazione la misura subordinate/AS-unit. Il grafico 3 e la tabella 12 mostrano come queste due misure hanno andamenti diversi, in alcuni casi addirittura inversi.

**Grafico 3. CAF – Il confronto tra le misure di subordinazione e di lunghezza – Italiane e Apprendenti (valori di gruppo)**



**Tabella 12. Le misure di subordinazione e di lunghezza, variazione**

	<i>film</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
parole/clausole	19,87%	27,40%	8,73%	7,14%
subordinate/AS-unit	-30,36%	-23,46%	-50,86%	11,34%

Per quanto riguarda le misure qualitative di gamma, sono state mantenute tutte le misure riportate nella tabella 9.

Vediamo ora le misure per rilevare l'accuratezza dell'interlingua orale. In §2.2.2.2 abbiamo visto le principali impiegate in letteratura, in questa sezione illustro la misura scelta per questo studio e ne giustifico l'utilità.

Il rapporto tra il numero di unità di produzione senza errori e il numero totale delle unità prodotte è unanimemente considerato una misura globale utile (anche se non esaustiva) per rilevare l'accuratezza delle produzioni a livello avanzato come sono i soggetti coinvolti in questo studio. Pertanto è stata qui utilizzata pur nella piena consapevolezza – è ovvio – che i risultati di un'analisi di questo tipo permetteranno di fare solo considerazioni globali sull'accuratezza.

In sede di analisi sono stati considerati errori le inesattezze di tipo morfo-sintattico o lessicale, come in (38-41), mentre vengono conteggiate come corrette quelle strutture che, pur deviando dallo standard, risultano corrette nell'italiano parlato, come in (42):

(38) PAN: || un: negoziere ||

(39) PAN: || {mh .h} e # lo pane prende uomo #0\_2 >quel de< cantie:re ||

(40) SHI: || .hhh ha fatto: di tutto | così la polizia gli: # venisse a prend fere, ||

(41) VAL: || {e: tch cioè: h} lui prende dei sigari ||  
VAL: || lo offre anche a dei £bambini£ || #0\_5

(42) VAL: || e si vede:: l'ultima scena | che si ritrovano su un pezzo di legno | a galleggiare:: ||

Inoltre nel caso in cui lo studente corregge un proprio errore, questo non viene conteggiato:

(43) SHI: || e poi {ha arriva} è arrivato sull'isola:: ||

Infine, nell'identificare le AS-unit senza errori sono stati esclusi gli errori di pronuncia:

(44) PAN: || | poi loro due contrano | ||

Per verificare l'affidabilità di questi criteri di misurazione dell'accuratezza è stato effettuato un controllo sperimentale su una parte del corpus: i task intervista, racconto di un film e apertura delle telefonate realizzati da Pandita ed Elisa a t1 (1278 parole, 446 AS-unit). Il controllo ha coinvolto due parlanti nativi nel ruolo di valutatori, cui è stato richiesto di identificare le AS-unit senza errori. Si è proceduto poi al confronto delle due codifiche, che ha dato un buon grado di affidabilità, con un accordo tra i due valutatori pari al 98%.

Vediamo ora la misurazione della fluenza. Come abbiamo visto in §2.2.2.2, le diverse misure di fluenza proposte in letteratura cercano di rilevarne tre aspetti: la fluenza legata alla velocità e la fluenza legata alle pause, caratterizzate dalle variabili temporali, e la fluenza legata alle riparazioni definita dai fenomeni di esitazione. In questa sezione presento le misure scelte per questo studio e ne giustifico l'utilità.

Il conteggio delle misure di fluenza ha un grosso limite pratico, poiché richiede prima una trascrizione di altissima qualità e dunque estremamente impegnativa in termini di tempo. Ai fini di questa tesi, tra le misure proposte in letteratura (cfr. § 2.2.2.2, tabella 4), non ne sono state prese in considerazione tre: velocità d'eloquio, durata delle pause e lunghezza media tra due pause. Il corpus e gli strumenti a mia disposizione non hanno infatti permesso il conteggio della prima misura, la velocità dell'eloquio. La seconda è stata scartata dopo un controllo sperimentale su una parte del corpus: il task racconto di una storia realizzato in t3 da Catherine, Elisa e Valentina. Qui i risultati prodotti dai due ricercatori che si sono occupati delle trascrizioni ha mostrato come possa esserci un buon grado di affidabilità nell'individuazione delle pause e nella distinzione tra micropause e pause lunghe (con un grado di accordo dell'87,5%) ma risulti meno affidabile il conteggio della durata delle

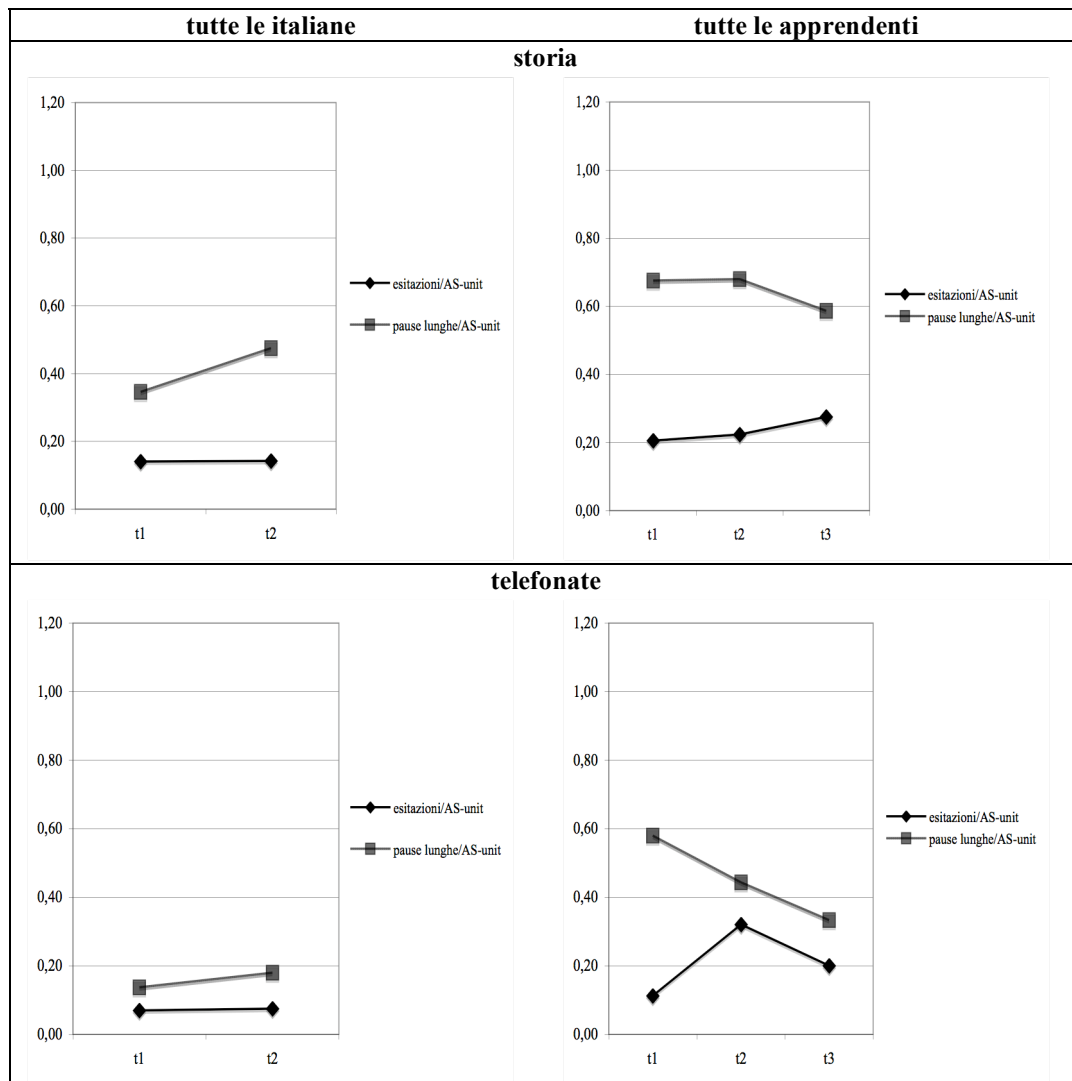


pause, con differenze anche molto marcate (accordo del 32% dei casi, con variazione media di #0\_3, e una deviazione standard di #0\_2/#0\_6). Si è pertanto ritenuto che il conteggio della durata media delle pause non potesse essere applicato ai dati a mia disposizione, in quanto avrebbe portato a risultati poco affidabili. Infine, l'esclusione della terza misura, lunghezza media tra due pause, è dovuta difficoltà di applicarla ai task interattivi, dove la definizione del confine dell'unità non è scontata. Poiché la conversazione è costruita insieme con l'interlocutore, il flusso del parlato si interrompe non solo nei casi in cui l'apprendente fa una pausa, piena o vuota, ma anche nel caso dei cambi di turno. Dal un lato considerare i cambi di turno come confine dell'unità significherebbe, nei task interattivi, ottenere unità molto brevi, dunque interpretabili come poco fluenti, dall'altro non considerarli confine di unità porterebbe a unità anche troppo lunghe, rendendo pertanto difficile l'interpretazione dei risultati.

In questo studio sono state scelte due misure di fluenza: numero di pause lunghe per AS-unit e numero di fenomeni di esitazione per AS-unit. Così come per la complessità, è stato effettuato un confronto tra le due misure su due task, il racconto di una storia ad immagini e l'apertura delle telefonate, realizzati dalle apprendenti e dalle italiane tra t1 e t3.

Il grafico 4 riporta i risultati ottenuti dai due gruppi sulle due misure di fluenza considerate e la tabella 13 ne riporta la variazione. Il confronto degli andamenti sui tre anni mostra come i risultati ottenuti siano piuttosto diversi l'uno dall'altro, indicando in alcuni casi tendenze addirittura opposte. Questo dato supporta l'idea suggerita in letteratura secondo la quale le misure considerate rilevano diversi e complementari aspetti della fluenza. Entrambe verranno pertanto impiegate per l'analisi del corpus oggetto di questo studio.

**Grafico 4. CAF – Il confronto tra le misure di fluenza pause lunghe e esitazioni – Italiane e apprendenti (valori di gruppo)**



**Tabella 13. Le misure di fluenza, variazione**

	<i>storia</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
pause lunghe/AS-unit	-13,20%	37,25%	-42,77%	50,45%
esitazioni/AS-unit	33,99%	1,31%	-11,63%	20,46%

Ricapitolando, in questo studio, riguardo al CAF vengono usate le seguenti misure: per la complessità sintattica due specifiche di grado (lunghezza e subordinazione); tre qualitative di gamma relative al tipo di clausola (percentuale di clausole indipendenti, unità indipendenti inferiori alla clausola e clausole subordinate sul totale delle clausole), e infine tre qualitative di gamma relative alla subordinazione (percentuale di avverbiali, argomentali e relative sul totale delle subordinate).

Per l'accuratezza viene usata la misura percentuale delle AS-unit senza errori sul totale delle AS-unit.

Infine per la fluenza sono state scelte due misure: il numero di pause lunghe per AS-unit, e quello di fenomeni di esitazione per AS-unit.

### 3.5.2. Le valutazioni QCER

In questa sezione presento le griglie del QCER utilizzate in questo studio, discuto un'indagine di validazione empirica dell'affidabilità delle valutazioni e infine illustro la procedura impiegata per la valutazione del mio corpus.

I descrittori selezionati per questo studio sono tratti dalle griglie relative agli aspetti qualitativi dell'uso della lingua parlata e si riferiscono alle tre dimensioni della coerenza-coesione, della correttezza grammaticale, e della fluenza. Questa selezione è dovuta al fatto che questi descrittori sono confrontabili con ciò che viene misurato attraverso le misure analitiche CAF esposte in § 3.5.1. Infatti alla complessità del CAF corrisponde la coerenza-coesione (QCER, p. 154), all'accuratezza corrisponde la correttezza grammaticale (QCER, p. 140), mentre – ovviamente – la fluenza del CAF coincide con quella del QCER (p. 158). La tabella 14 riporta i descrittori del QCER selezionati.

**Tabella 14. QCER – Gli aspetti qualitativi dell’uso della lingua parlata**

	<i>coerenza-coesione</i>	<i>correttezza grammaticale</i>	<i>fluenza</i>
<b>C2</b>	È in grado di realizzare un discorso coerente e coeso usando in modo appropriato una grande varietà di schemi organizzativi e un’ampia gamma di connettivi e di meccanismi coesivi di altro tipo.	Mantiene costantemente il controllo grammaticale di forme linguistiche complesse, anche quando la sua attenzione è rivolta altrove (ad es. nella pianificazione di quanto intende dire e nell’osservazione delle reazioni altrui).	È in grado di esprimersi in discorsi lunghi con scioltezza naturale, senza sforzi né esitazioni. Fa pause solo per cercare le parole per esprimere con precisione il proprio pensiero o per trovare un esempio o una spiegazione adatti.
<b>C1</b>	È in grado di realizzare un discorso chiaro, sciolto e ben strutturato, mostrando un uso controllato degli schemi organizzativi, di connettivi ed espressioni coesive.	Mantiene costantemente un livello elevato di correttezza grammaticale; gli errori sono rari e poco evidenti.	È in grado di esprimersi con scioltezza e spontaneità, quasi senza sforzo. Solo un argomento concettualmente complesso può ostacolare la naturale scioltezza del discorso.
<b>B2+</b>	È in grado di usare in modo efficace diversi connettivi per esplicitare i rapporti tra i concetti.	Ha una buona padronanza grammaticale; nella struttura delle frasi possono ancora verificarsi sbagli occasionali, errori non sistematici e difetti minori, che sono per lo più rari e corretti a posteriori.	È in grado di comunicare con spontaneità, dando per lo più prova di notevole scioltezza e uso disinvolto dei mezzi espressivi anche in discorsi piuttosto lunghi e complessi.
<b>B2</b>	È in grado di usare un numero limitato di elementi di coesione per collegare i propri enunciati in un discorso chiaro e coerente. In un intervento lungo possono presentarsi dei “salti” logici.	Mostra una padronanza grammaticale piuttosto buona. Non fa errori che possano provocare fraintendimenti.	È in grado di produrre sequenze discorsive con un ritmo abbastanza uniforme; anche se può avere delle esitazioni quando cerca strutture ed espressioni, fa poche pause evidenti. È in grado di interagire con spontaneità e scioltezza tali da consentire una normale interazione con parlanti nativi, senza sforzi per entrambe le parti.
<b>B1+</b>		Comunica con ragionevole correttezza in contesti familiari; la padronanza grammaticale è generalmente buona anche se si nota l’influenza della lingua madre. Nonostante gli errori, ciò che cerca di esprimere è chiaro.	È in grado di esprimersi con relativa disinvoltura. Nonostante alcuni problemi di formulazione che possono sfociare in pause e blocchi, è in grado di portare avanti il discorso efficacemente senza aiuto.

<b>B1</b>	È in grado di collegare una serie di elementi relativamente brevi e semplici in una sequenza lineare per punti.	Usa in modo ragionevolmente corretto un repertorio di formule di <i>routine</i> e strutture d'uso frequente, relative alle situazioni più prevedibili.	È in grado di mantenere il discorso comprensibile, anche se sono evidenti pause per cercare parole e forme grammaticali e per riparare agli errori, specialmente nelle sequenze di produzione libera di una certa lunghezza.
<b>A2+</b>	È in grado di collegare frasi semplici usando i connettivi più usuali per raccontare una storia o descrivere qualcosa, realizzando un semplice elenco di punti.		Riesce a farsi comprendere con enunciati molto brevi, nonostante che pause, false partenze e riformulazioni siano molto evidenti.
<b>A2</b>	È in grado di collegare gruppi di parole con connettivi semplici quali "e", "ma" e "perché".	Usa correttamente alcune strutture semplici, ma continua sistematicamente a fare errori di base – per esempio tende a confondere i tempi verbali e a dimenticare di segnalare gli accordi; ciononostante ciò che cerca di dire è solitamente chiaro.	Su argomenti familiari è in grado di formulare espressioni e ha sufficiente capacità di portare a termine scambi comunicativi brevi, nonostante le esitazioni e le false partenze siano evidenti.
<b>A1</b>	È in grado di collegare parole o gruppi di parole con connettivi molto elementari quali "e" o "allora".	Ha solo una padronanza limitata di qualche struttura grammaticale e di semplici modelli sintattici, in un repertorio memorizzato.	È in grado di cavarsela con enunciati molto brevi, isolati, solitamente memorizzati, facendo molte pause per cercare le espressioni, per pronunciare le parole meno familiari e per riparare agli errori di comunicazione.

Per rilevare l'affidabilità nell'applicazione dei descrittori selezionati, mi rifaccio all'indagine sperimentale condotta da Ferrari & Nuzzo (in stampa). Tale indagine consiste in una validazione empirica dell'affidabilità delle valutazioni espresse da due gruppi di valutatori, tutti insegnanti, tra cui 5 esperti valutatori e 5 non esperti, su tre attività del corpus oggetto di questo studio (il racconto di un film, l'intervista e le telefonate) realizzate da 3 soggetti (due apprendenti, Pandita e Shirley, e una studentessa italiana, Elisa). Gli insegnanti esperti sono docenti di un Centro Territoriale Permanente, formati come esaminatori e regolarmente coinvolti come valutatori della prova

orale negli esami di certificazione CELI dell'Università per Stranieri di Perugia. Gli insegnanti non esperti hanno diverse esperienze nel mondo della scuola, e sono tutti impegnati come volontari in scuole di italiano L2 per adulti.

Ciascun gruppo di insegnanti è stato coinvolto in una sessione di valutazione della durata di 3 ore circa durante la quale si sono valutate le videoregistrazioni dei 3 task, riferendosi a griglie basate sia sui descrittori degli aspetti qualitativi dell'uso della lingua parlata che sui descrittori specifici riferiti al task. L'attribuzione del livello avveniva subito dopo la visione di ciascuna produzione ed era suddivisa in due momenti: una prima valutazione individuale e una seconda valutazione condivisa da raggiungere attraverso una discussione di gruppo. L'analisi dell'affidabilità dei giudizi espressi (statistica descrittiva e correlazione) è stata condotta sulle valutazioni individuali e sulle valutazioni condivise.

La tabella 15 riporta i risultati relativi ai tre descrittori impiegati per questo studio. Ricordando che in letteratura in genere si considerano affidabili i giudizi quando il valore della correlazione è pari o superiore a 0,80 (Alderson *et al.*, 1995), osserviamo che i descrittori relativi alla correttezza grammaticale hanno ottenuto il maggior accordo. Rispetto a coerenza-coesione le discrepanze più rilevanti sono relative al task telefonate, poiché in diversi casi le insegnanti non sono state in grado di esprimere una valutazione ritenendo i descrittori inadeguati per un task così interattivo. I risultati relativi alla fluenza sono stati invece meno soddisfacenti.

**Tabella 15. L'affidabilità dei giudizi. Spearman rank-order  
(da Ferrari & Nuzzo, in stampa)**

	<i>tutti gli insegnanti</i>	<i>insegnanti esperti</i>	<i>insegnanti non esperti</i>	<i>giudizi comuni</i>
coerenza-coesione	0,83	0,89	0,83	0,79
correttezza grammaticale	0,90	0,88	0,81	0,87
fluenza nel parlato	0,68	0,87	0,55	0,83

Gli insegnanti esperti dimostrano maggior accordo rispetto ai colleghi non esperti. Questo risultato è in linea con ciò che viene da più parti

evidenziato nella letteratura sul testing, e cioè che la formazione aumenta la coerenza e l'affidabilità delle scelte dei valutatori (Lunz *et al.*, 1990) oltre che ridurre le differenze di severità tra esaminatori (Wigglesworth, 1993; Weigle, 1994). Lumley & McNamara (1995) suggeriscono che i vantaggi della formazione non sono però di lunga durata, dimostrando come sia necessario un momento di formazione prima di ogni sessione di valutazione. Interessanti inoltre i risultati ottenuti nelle valutazioni condivise: la discussione nei gruppi per il raggiungimento di un accordo sembra portare gli insegnanti ad allinearsi maggiormente nelle loro scelte. Questo dato sostiene l'utilità della pratica diffusa nelle certificazioni di impiegare due valutatori per ciascuna valutazione (cfr. Alderson *et al.*, 1995). Nonostante non ci sia sempre una correlazione ottimale dei giudizi per i 10 insegnanti, i risultati di questa indagine permettono alcune osservazioni in positivo. Sugli 11 livelli disponibili<sup>2</sup> per ciascuna competenza, lo scarto tra le scelte effettuate oscilla tra 2 e 4 livelli, con una preferenza per 2 valutazioni diverse. Le diverse valutazioni inoltre si riferiscono nella maggior parte dei casi a livelli contigui sulla scala. Questo suggerisce come il QCER possa essere un buon punto di riferimento, anche se sembra non riuscire a guidare in modo chiaro e preciso le scelte degli insegnanti nella distinzione tra livelli contigui, in particolare per la fascia C. Del resto difficoltà per i livelli più alti sono state evidenziate anche in altre ricerche: Kaftandjieva & Takala (2002: 113), ad esempio, sottolineano come il livello C2 non sia sempre distinto dal C1; Cassandro & Maggini (2004: 71), conducendo un'esercitazione con un gruppo di insegnanti, osservano come non manchino divergenze vistose nell'attribuzione dei livelli, specialmente per quanto riguarda la parte alta del continuum. Questo fa presupporre che anche se la formazione appare come un elemento centrale nel garantire l'affidabilità delle valutazioni, non si possono escludere fattori interni al QCER per la scarsa affidabilità dei giudizi, dal momento che anche tra i valutatori esperti si riscontrano problemi di allineamento.

---

<sup>2</sup> Calcolando livelli e sotto-livelli, dal momento che in molti casi i livelli delle fasce A e B sono ulteriormente suddivisi (per esempio, A1, A1+, A2, A2+).

Poiché insegnanti esperti e valutazioni condivise ottengono i risultati migliori, nel definire la procedura di valutazione per questo studio si è cercato di soddisfare entrambe le condizioni. Sono stati coinvolti due valutatori esperti, entrambi regolarmente impegnati come esaminatori in esami di certificazione, oltre che buoni conoscitori del QCER. Ai valutatori è stato richiesto di esprimere un giudizio su complessità, accuratezza e fluenza delle produzioni di tutte le apprendenti del corpus in tre task (intervista, storia e telefonate) realizzati a t1 e a t4, utilizzando le griglie del QCER predisposte. Sono state così valutate in 3 sessioni di 2 ore ciascuna 36 videoregistrazioni, randomizzate per task, apprendenti e momento di rilevazione. Ai valutatori è stato richiesto di esprimere prima una valutazione individuale e di concordare poi una valutazione condivisa. I risultati della valutazione individuale sono stati utilizzati per verificare l'*inter-rater reliability* dei due valutatori. I risultati sono stati piuttosto buoni: 0,82 per coerenza-coesione, 0,88 per correttezza grammaticale, e 0,84 per fluenza. I risultati della valutazione condivisa sono stati poi utilizzati per osservare la variazione longitudinale e situazionale di complessità, accuratezza e fluenza, oggetto di indagine di questo studio.





## **CAPITOLO 4**

### **RISULTATI**

Il capitolo è dedicato alla presentazione dei risultati delle analisi effettuate sul corpus. Prima verranno presentati i risultati dell'analisi CAF (§§ 4.1-4.4), poi quelli delle valutazioni QCER (§ 4.5). Infine verranno discusse alcune interazioni tra i due approcci (§ 4.6).

#### *4.1. La complessità sintattica CAF*

La sezione presenta i risultati relativi all'analisi della complessità sintattica effettuata attraverso l'applicazione delle due misure globali: lunghezza (parole/clausole) e subordinazione (subordinate/AS-unit); e delle sei misure specifiche: tipo di clausola (percentuale di clausole indipendenti, di unità indipendenti inferiori alla clausola e di subordinate sul totale delle clausole) e tipo di subordinata (percentuale di avverbiali, di argomentali e di relative sul totale delle subordinate).

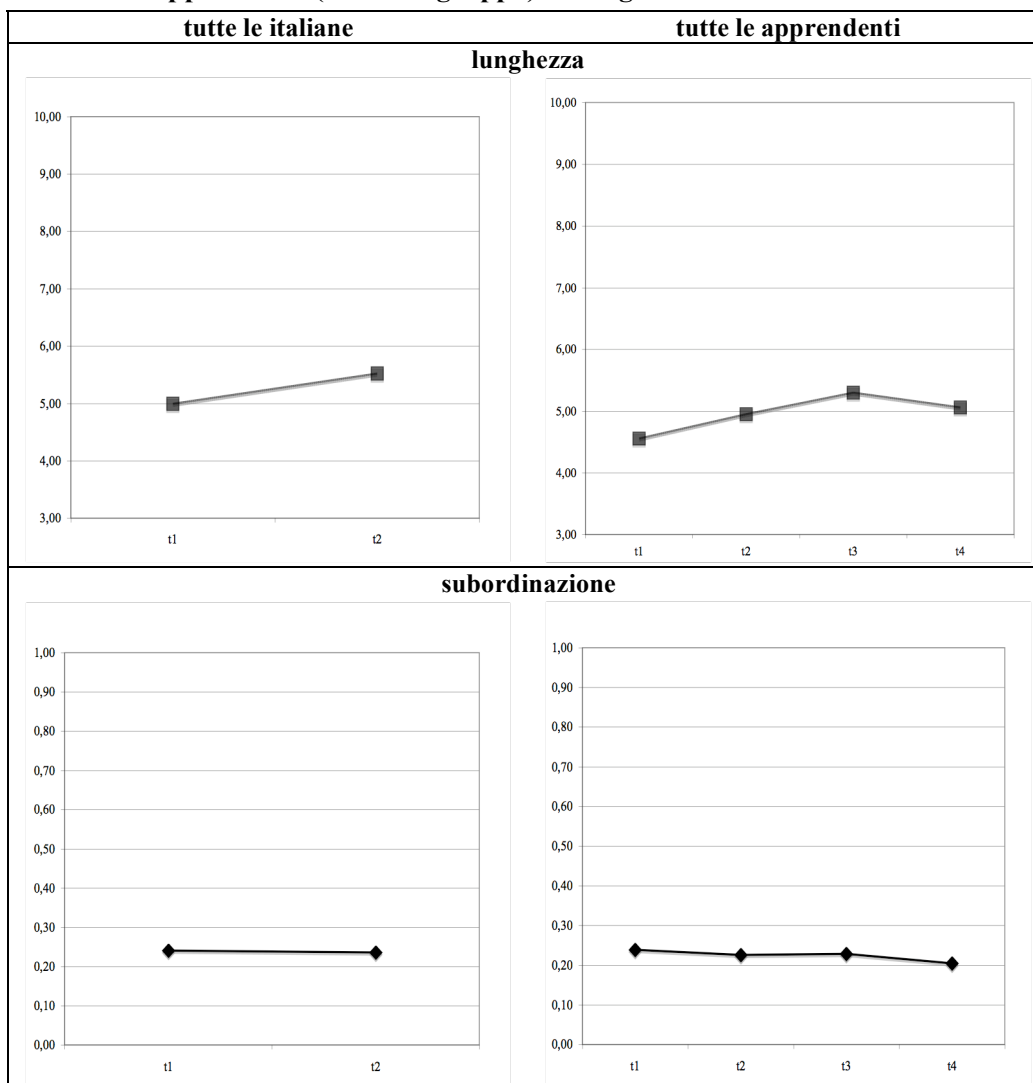
Si inizia ogni volta con l'analisi della variazione longitudinale della complessità sintattica considerando complessivamente i due gruppi di soggetti (studentesse italiane e apprendenti) cumulativamente in tutti i task (§ 4.1.1); si prosegue con quella della variazione situazionale individuando le differenze tra i quattro task (§ 4.1.2), e si conclude illustrando, con esempi oltre che con grafici, il percorso individuale di ogni studentessa (§ 4.1.3).

#### 4.1.1. La variazione longitudinale

Il grafico 5 mostra i valori delle due misure di lunghezza e di subordinazione relativamente allo sviluppo longitudinale, mentre la tabella 16 riporta quelli delle sei misure relative al tipo di clausola e di subordinata. In entrambi i casi i dati si riferiscono ai valori medi dei due gruppi considerati cumulativamente per tutti i task.

I dati mostrano come già a t1 le apprendenti sono piuttosto avanzate rispetto alla complessità sintattica. Infatti la misura di subordinazione ottiene gli stessi valori delle compagne italiane, e la misura di lunghezza è di poco inferiore. Alcune differenze si registrano però rispetto alla distribuzione qualitativa di tipi di clausole e di subordinate, poichè le apprendenti producono più unità indipendenti inferiori alla clausola rispetto alle italiane, e utilizzano subordinate relative e argomentali piuttosto che avverbiali. In prospettiva longitudinale si realizzano due tipi di progresso. In primo luogo, rispetto alla lunghezza vi è un aumento dei valori, e in secondo luogo, nonostante per la subordinazione ci sia una lievissima flessione della misura, la distribuzione qualitativa del tipo di clausole e di subordinate si avvicina man mano a quella nativa.

**Grafico 5. CAF – Complessità – Variazione longitudinale – Tutti i task – Italiane e apprendenti (valori di gruppo) – Lunghezza e subordinazione**



**Tabella 16. CAF – Complessità – Variazione longitudinale – Tutti i task – Italiane e apprendenti – (valori di gruppo) – Tipi di clausole e di subordinate**

	<i>t1</i>	<i>t2</i>	<i>t1</i>	<i>t2</i>	<i>t3</i>	<i>t4</i>
	<i>italiane</i>		<i>apprendenti</i>			
clausole indep.	61,89%	61,62%	58,30%	61,00%	60,88%	64,39%
unità inf. clausola	19,55%	19,56%	23,58%	21,54%	22,59%	20,76%
clausole sub.	18,56%	18,82%	18,12%	18,62%	17,94%	16,15%
avverbiali/sub.	28,84%	44,44%	42,97%	42,78%	48,23%	54,28%
argomentali/sub.	44,40%	30,27%	28,47%	32,51%	28,60%	24,42%
relative/sub.	26,77%	40,11%	33,74%	26,96%	32,27%	25,45%

Questo dato è in linea con quanto suggerito in letteratura con l'Ipotesi di Predizione dello Sviluppo proposta recentemente da Wolfe-Quintero *et al.* (1998; cfr. § 2.2.2.2), secondo la quale ai livelli più avanzati lo sviluppo della complessità non riguarda tanto l'aumento della misura di subordinazione, bensì la complessificazione a livello di clausola, con un conseguente aumento della lunghezza media delle clausole, accompagnato nei miei dati, anche da un uso più vario di tipi di subordinate. Questi risultati giustificano l'utilità della combinazione dei tre tipi di misure nell'analisi della complessità sintattica a livello avanzato.

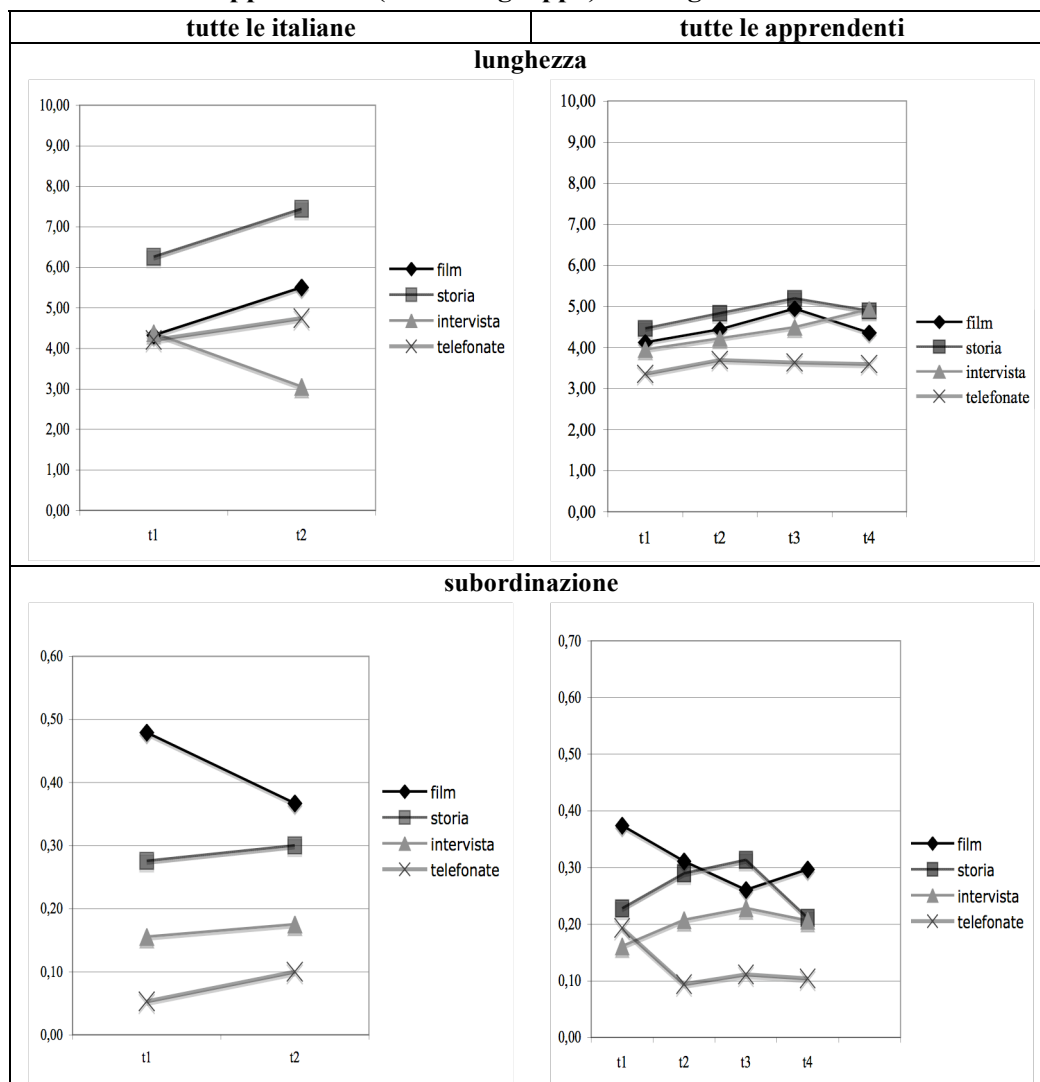
#### 4.1.2. La variazione situazionale

Il grafico 6 mostra i valori della complessità sintattica delle produzioni dei due gruppi di soggetti relativi alle misure di lunghezza e di subordinazione. La tabella 17 riporta invece i dati relativi alle sei misure del tipo di clausola e di subordinata riscontrati nei quattro task. In entrambi i casi vengono forniti, task per task, i valori medi dei due gruppi.

I dati disaggregati per task offrono un quadro in parte diverso da quello descritto con la sola longitudinalità in § 4.1.1. Vediamo nel dettaglio la variazione situazionale segnata da italiane e apprendenti.

Le studentesse italiane dimostrano un certo grado di variazione tra task: i task monologici film e storia stimolano la produzione di clausole più lunghe, e di AS-unit più complesse, con un uso di diversi tipi di clausole e di subordinate, quest'ultime soprattutto avverbiali, mentre i task più interattivi intervista e telefonate fanno realizzare AS-unit più semplici e brevi, con un maggior uso di unità indipendenti inferiori alla clausola e un uso ridotto di subordinate, soprattutto argomentali o relative. Questo dato è in accordo con gli studi di approccio psicolinguistico (§ 2.1.2), secondo cui i diversi task pongono diverse richieste ai parlanti in termini di complessità della lingua

**Grafico 6. CAF – Complessità – Variazione situazionale – Distribuzione per task  
– Italiane e apprendenti (valori di gruppo) – Lunghezza e subordinazione**



**Tabella 17. Complessità – Variazione situazionale – Distribuzione per task – Italiane e apprendenti – (valori di gruppo) – Tipi di clausole e di subordinate<sup>1</sup>**

	<i>t1</i>	<i>t2</i>	<i>t1</i>	<i>t2</i>	<i>t3</i>	<i>t4</i>
	<i>italiane</i>		<i>apprendenti</i>			
<i>film</i>						
clausole indep.	58,83%	56,63%	63,91%	65,93%	68,07%	66,33%
unità inf. clausola	10,05%	16,35%	9,45%	10,52%	11,37%	10,93%
clausole sub.	31,12%	27,04%	26,65%	23,56%	20,57%	22,74%
avverbiali/sub.	20,19%	28,57%	39,41%	44,37%	59,03%	39,79%
argomentali/sub.	50,97%	36,54%	29,70%	39,41%	24,80%	27,78%
relative/sub.	28,85%	49,18%	30,89%	21,63%	49,50%	32,43%
<i>storia</i>						
clausole indep.	68,25%	72,03%	68,17%	70,26%	66,78%	71,71%
unità inf. clausola	10,69%	4,89%	13,60%	8,04%	12,95%	11,25%
clausole sub.	21,07%	23,09%	18,23%	21,70%	23,51%	17,04%
avverbiali/sub.	35,55%	34,71%	27,34%	35,31%	35,09%	43,11%
argomentali/sub.	38,00%	32,65%	30,43%	27,44%	34,28%	25,19%
relative/sub.	26,46%	32,65%	42,24%	37,25%	30,64%	31,71%
<i>intervista</i>						
clausole indep.	60,54%	51,86%	53,72%	52,42%	56,31%	64,47%
unità inf. clausola	26,12%	32,99%	32,65%	30,50%	25,12%	26,13%
clausole sub.	13,35%	15,16%	13,64%	17,08%	18,58%	15,93%
avverbiali/sub.	30,77%	44,73%	62,16%	48,67%	53,27%	79,96%
argomentali/sub.	44,23%	41,22%	24,21%	30,67%	25,76%	18,93%
relative/sub.	25,00%	14,06%	26,23%	20,67%	20,97%	7,79%
<i>telefonate</i>						
clausole indep.	59,95%	64,42%	47,41%	55,39%	52,37%	55,04%
unità inf. clausola	31,36%	21,66%	38,62%	37,10%	38,50%	36,08%
clausole sub.	8,70%	13,93%	13,97%	10,02%	9,13%	8,89%
avverbiali/sub.	-	-	x	x	x	x
argomentali/sub.	x	x	x	x	x	x
relative/sub.	x	x	x	x	x	x

<sup>1</sup> I valori riportati nella tabella si riferiscono ai valori medi delle diverse telefonate del corpus. Poiché questo task non richiede una produzione linguistica consistente, i dati qui forniti devono essere con una certa cautela. Viste poi le scarse occorrenze di subordinate nel corpus, indicare la percentuale dei tipi di subordinate sarebbe stato poco significativo e avrebbe dato una lettura falsata dei risultati, per questo nella tabella si è indicata esclusivamente la presenza (x) o meno (-) della struttura.

prodotta. Nella ricostruzione di un racconto il flusso della conversazione è gestito quasi esclusivamente dal soggetto e l'intervistatore si limita a segnali di assenso. Quindi la narrazione procede con frasi piuttosto lunghe e maggiore subordinazione. Nei due task interattivi, seppure a livelli diversi, la conversazione richiede invece una maggior collaborazione tra soggetto e interlocutore, favorendo quindi una maggior alternanza di turni e l'uso di AS-unit più brevi e semplici.

Longitudinalmente, le studentesse italiane mostrano alcuni cambiamenti. Nei task monologici si riduce l'uso di subordinate a favore di clausole un po' più lunghe, mentre nei task interattivi non si verifica alcun cambiamento sostanziale, se non una riduzione in lunghezza per l'intervista e un lieve aumento in complessità per le telefonate. Vedremo in § 4.1.3 come per le telefonate l'incremento sia legato all'uso di formule più complesse di richiesta.

Vediamo ora i dati relativi alle apprendenti. A t1, differenziandosi dalle compagne italiane, dimostrano non solo molta meno variazione tra task, ma anche una cruciale differenza nella direzione di tale variazione. Infatti l'apertura delle telefonate ottiene valori di subordinazione pari a quelli delle due narrazioni. Se nei task monologici e nell'intervista le differenze rispetto alle italiane sono minime, in quello di apertura delle telefonate la distanza dalle compagne italiane è piuttosto evidente. Vedremo in dettaglio attraverso estratti del corpus in § 4.1.3 come le apprendenti in questo task producono sia un maggior numero che una maggior varietà di tipi di subordinate rispetto alle compagne italiane: esse tendono a condensare diverse informazioni in una stessa unità sintattica piuttosto che gestire gli scambi della routine dell'apertura di una telefonata di servizio con enunciati rapidi e brevi. In altre parole, le apprendenti dimostrano di aver già acquisito la capacità di produrre costruzioni complesse, sanno realizzare subordinate e riescono a condensare un alto numero di parole in un unico progetto sintattico. Tuttavia, esse paiono meno in grado di scegliere quando utilizzare tali costruzioni. Saper gestire l'apertura di una telefonata non implica solamente il controllo delle regole sintattiche, ma anche la capacità di saper produrre in rapida successione turni che siano



contingenti rispetto alle produzioni altrui. Questo tipo di abilità è ancora carente a t1 per le nostre apprendenti, per quanto di livello relativamente avanzato.

In prospettiva longitudinale, le apprendenti realizzano diversi progressi differenziando come le native tra task: la misura di lunghezza registra lievi aumenti con traiettorie lineari in tutti e quattro i task, confermando dunque il quadro definito dai valori di gruppo, mentre la misura di subordinazione e la distribuzione del tipo di subordinate varia, seguendo traiettorie che si differenziano per intensità e direzione da task a task, ma ottenendo a t4 in tutti i casi un avvicinamento al quadro di variazione definito per le italiane. Sebbene le apprendenti differenzino meno delle compagne italiane, all'ultima rilevazione lo fanno allo stesso modo. Vediamo in dettaglio la variazione segnata da ogni task.

Così come per le italiane, nei task monologici la traiettoria di sviluppo non è lineare, nel film la misura di subordinazione decresce tra t1 e t3, per poi aumentare nuovamente, mentre nella storia l'andamento è opposto. I cambiamenti quantitativi sono accompagnati, nel task storia da avvicinamenti qualitativi della distribuzione dei tipi di clausole e di subordinate, nel task film la riduzione in subordinazione è invece controbilanciata da un aumento della lunghezza delle clausole prodotte.

Diverso l'andamento nei task interattivi. Nel task intervista lo sviluppo è più lineare, le apprendenti dimostrano un generale aumento tra t1 e t4, sia della lunghezza che della subordinazione, oltre che un uso più vario di tipi di subordinate. È bene sottolineare come in questo task le apprendenti parlino addirittura di più rispetto alle compagne italiane, dando più informazioni e richiedendo un minor numero di domande o interventi da parte dell'intervistatore.

E infine, nel task telefonate l'avvicinamento al quadro di variazione nativo è ottenuto attraverso una riduzione lineare della misura di subordinazione e dei tipi di subordinate impiegate.

Ricapitolando, in § 4.1.1 abbiamo visto come la complessità misurata in termini di subordinazione possa svilupparsi secondo percorsi non lineari a livelli avanzati (cfr. Ipotesi di Predizione dello Sviluppo, § 2.2.2.2). I dati disaggregati per task confermano questa tendenza nei task monologici, ma non in quelli interattivi. Questo risultato suggerisce quindi una seconda possibile ragione dell'andamento non lineare di questa misura: il tipo di attività svolta. Come abbiamo visto, nel task interattivo l'alto valore di subordinazione evidenzia incompetenza sul piano interazionale: trovandosi a gestire uno scambio rapido e contingente come quello richiesto dall'apertura di una telefonata, le nostre apprendenti, ansiose di dire subito tutto il dovuto, a t1 preferiscono condensare le informazioni in unità complesse, che probabilmente reputano di potere gestire meglio, piuttosto che realizzare turni brevi e incalzanti rispetto a quelli dell'interlocutore. A t2 e t3 il controllo più nativo del task è ottenuto di conseguenza attraverso una riduzione della complessità sintattica. Il processo di apprendimento di una lingua non è infatti caratterizzato semplicemente dalla crescita del repertorio sintattico dell'apprendente, ma anche dall'abilità ad utilizzarlo appropriatamente in una varietà di situazioni.

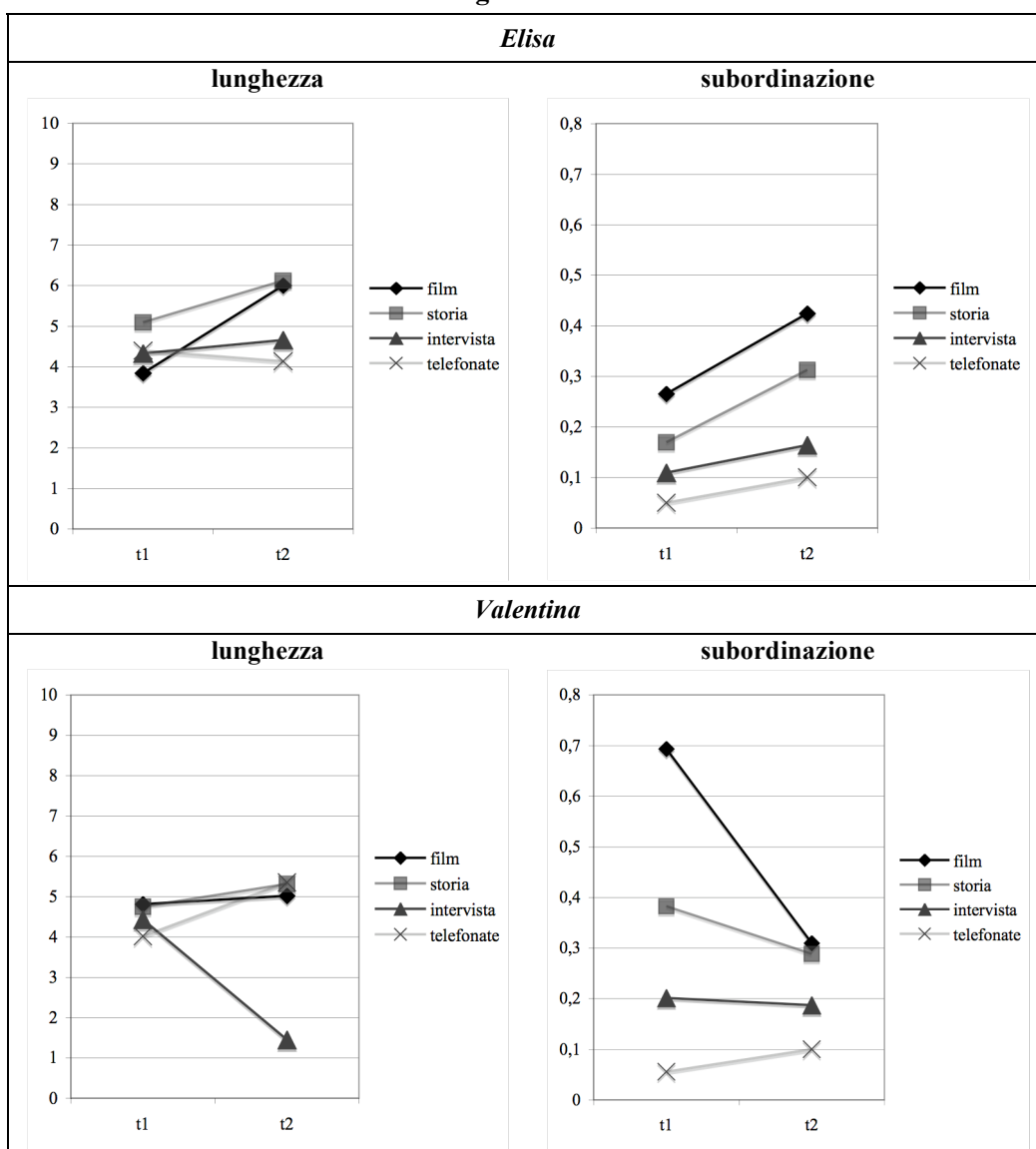
#### 4.1.3. La variazione individuale

I grafici 7 e 8 mostrano soggetto per soggetto i valori della complessità relativi alle misure di lunghezza e di subordinazione, e le tabelle 18-23 quelli relativi alle sei misure del tipo di clausola e di subordinata. La discussione delle cifre è accompagnata dall'illustrazione di esempi dal corpus.

Il grafico 7 e le tabelle 18 e 19 riportano i valori individuali ottenuti dalle due studentesse italiane. Pur confermando il quadro di variazione identificato in § 4.1.2, si registra variazione individuale soprattutto nei task monologici. In questi task alla prima rilevazione Elisa sviluppa narrazioni brevi e semplici, organizzate attraverso la giustapposizione di clausole indipendenti, Valentina,

al contrario, produce racconti già lunghi e sintatticamente complessi, in particolare nella narrazione del task relativo al film. Le differenze quantitative sono accompagnate da differenze qualitative della distribuzione di tipi di clausole e di subordinate. In prospettiva longitudinale si riduce la variazione individuale.

**Grafico 7. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Italiane – Lunghezza e subordinazione**



**Tabella 18. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Elisa – Tipi di clausole e di subordinate**

<i>t1</i>					<i>t2</i>			
<i>film</i>					<i>storia</i>			
clausole indep.	39	62,90%	27	58,70%	59	77,63%	29	69,05%
unità inf. clausola	10	16,13%	5	10,87%	6	7,89%	3	7,14%
clausole sub.	13	20,97%	14	30,43%	11	14,47%	10	23,81%
avverbiali/sub.	1	7,69%	4	28,57%	5	45,45%	4	40,00%
argomentali/sub.	7	53,85%	7	50,00%	3	27,27%	3	30,00%
relative/sub.	5	38,46%	3	21,43%	3	27,27%	3	30,00%
<i>intervista</i>					<i>telefonate</i>			
clausole indep.	80	61,07%	40	56,34%	3,25	60,41%	4,00	59,26%
unità inf. clausola	38	29,01%	21	29,58%	1,63	30,30%	1,75	25,93%
clausole sub.	13	9,92%	10	14,08%	0,50	9,29%	1,00	14,81%
avverbiali/sub.	3	23,08%	3	30,00%	-	-	-	-
argomentali/sub.	5	38,46%	5	50,00%	x	-	x	-
relative/sub.	5	38,46%	2	20,00%	-	-	x	-

**Tabella 19. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Valentina – Tipi di clausole e di subordinate**

<i>t1</i>					<i>t2</i>			
<i>film</i>					<i>storia</i>			
clausole indep.	69	54,76%	30	54,55%	83	58,87%	57	75,00%
unità inf. clausola	5	3,97%	12	21,82%	19	13,48%	2	2,63%
clausole sub.	52	41,27%	13	23,64%	39	27,66%	17	22,37%
avverbiali/sub.	17	32,69%	-	-	10	25,64%	5	29,41%
argomentali/sub.	25	48,08%	3	23,08%	19	48,72%	6	35,29%
relative/sub.	10	19,23%	10	76,92%	10	25,64%	6	35,29%
<i>intervista</i>					<i>telefonate</i>			
clausole indep.	93	60,00%	108	47,37%	3,67	59,48%	4,00	69,57%
unità inf. clausola	36	23,23%	83	36,40%	2,00	32,41%	1,00	17,39%
clausole sub.	26	16,77%	37	16,23%	0,50	8,10%	0,75	13,04%
avverbiali/sub.	10	38,46%	22	59,46%	-	-	-	-
argomentali/sub.	13	50,00%	12	32,43%	x	-	x	-
relative/sub.	3	11,54%	3	8,11%	x	-	x	-

Vediamo attraverso alcuni esempi come le italiane realizzano i quattro task e quali sono le tendenze della variazione individuale. Nei task monologici, alla prima rilevazione le due studentesse italiane utilizzano stili e strategie diverse. Elisa sviluppa una narrazione sintetica, riporta solo i punti principali della trama attraverso la giustapposizione o la coordinazione di AS-unit brevi e semplici, così come illustrato dall'esempio (45) relativo al task film a t1.

- (45) ELI: || | {eh. che dice che:: # lui e: tch vedo che} cioè ho visto | che va:: {a trovar un va} a cercare un lavoro? | ||  
 INV: mhm #1\_4  
 ELI: || | e però va a finir male. | || #  
 ELI: || | fa::: un casino | ||  
 ELI: || | diciamo | ||  
 ELI: || | c'è una nave | {che:: fhhh}   
 INV: fhhh  
 ELI: .hh fhh che:: se ne va | ||  
 ELI: || | diciamo: | ||  
 ELI: || | lui s- muove un: e::: | || #1\_6  
 ELI: || | non lo so. | ||  
 ELI: || | comunque lui lavora in un cantiere tipo | ||  
 INV: mhm  
 ELI: || | insomma alla fine lo cacciano via? | || #0\_6  
 ELI: || | poi::: per la strada incontra questa ragazza {che è:: #0\_7 } | ||  
 ELI: || | {non ha: mh} cioè è una ragazza povera:: | ||  
 ELI: || | non ha da mangiare:: | ||  
 ELI: || | è da sola:: | ||  
 ELI: || | così. | ||  
 ELI: || | che:: ruba:: un filoncino di pane::? | || #0\_4

Qui Elisa racconta una delle scene iniziali del video: la studentessa riporta informazioni generiche sugli eventi, senza scendere nei dettagli. Realizza due subordinate, la prima argomentale retta da *vedere* e la seconda realizzata da una relativa introdotta dal presentativo *c'è*, non certo considerabili esempi di subordinazione sofisticata. La lunghezza media delle clausole è di 4,5 parole. L'utilizzo di segnali discorsivi come *comunque*, *insomma* e *diciamo* o l'uso del *che* sintatticamente inattivo compensano però la brevità delle clausole e l'uso ridotto di subordinazione, creando un effetto di dipendenza semantica tra gli enunciati.

Diversamente dalla compagna, Valentina realizza i due task monologici producendo resoconti lunghi il doppio. La sua narrazione è piuttosto dettagliata, caratterizzata da un uso importante della subordinazione come

evidente dall'estratto (46), realizzato nel task film a t1. Qui la studentessa riporta la stessa scena presa ad esempio per la compagna realizzando però una narrazione in cui le informazioni sono impacchettate entro AS-unit più lunghe e complesse.

- (46) VAL: || e::m tch come lavoro gli danno::: | quello di::: andare::: mh tch in una specie {di::: h #0\_5 cioè di me di::: h} di porto | dove ci son tutte delle navi::: #0\_4 così ||  
 VAL: || e c'è anche un uomo||  
 INV: in un cantiere  
 VAL: || {un cant-} sì||  
 INV: cantiere navale tipo  
 VAL: || sì un cantiere navale così || h h  
 VAL: || e c'è un uomo? | in pratica che:: gli dice | che deve iniziare a lavorare | per vedere | se {era} appunto::: h riusciva::: | a trovarsi bene||  
 VAL: || doveva portargli un cuneo || #0\_4  
 INV: m m  
 VAL: || poi #0\_4 subito::: gli prende un cuneo: lungo || #0\_3  
 VAL: || che in pratica {era::} il suo datore gli aveva detto | di portargliene uno: triangolare così||  
 INV: mh mh #0\_8  
 VAL: || hh e::: dice ||  
 VAL: || vallo a cercare || #0\_2  
 VAL: || e lui va a cercarlo || #0\_6  
 VAL: || ne prende uno ||  
 VAL: || solo che dopo £parte la nave£ he ?£h£ h || # 1\_0  
 VAL: || quindi in pratica ha fatto un guaio h || #0\_8  
 VAL: || e quindi::: l'hanno licenziato ||  
 INV: ?£h£ h? ah #0\_6  
 VAL: || tch h e::m dopo che::: l'hanno licenziato | in pratica: sta passeggiando lungo la strada h | quando vede: una ragazza #0\_2 | che: {sta ruba cioè} sta rubando un filoncino di pane ||

La codifica del tipo di clausole e l'analisi della loro frequenza confermano le tendenze osservare finora (tabelle 18-19). Nel racconto di Elisa, la ragazza che parla meno e con meno subordinate, viene preferito l'uso di clausole indipendenti (62,90%) e unità indipendenti inferiori alla clausola (16,13%) rispetto alle clausole subordinate, il cui uso è ridotto a poche occorrenze (20,97%), prevalentemente argomentali e relative. Abbiamo visto precedentemente come lo scarso uso di subordinate sia compensato dall'uso di clausole indipendenti introdotte dal connettore *che* sintatticamente inattivo e da segnali discorsivi. A questi si aggiunge l'uso frequente di argomentali con forme di subordinazione lassa, utile comunque a dare una certa coesione al racconto. Nell'esempio (47) l'indipendente *si vedono in questa casa* regge due

clausole introdotte dal complementatore generico *che*: *che stanno bene* e *che c'è la mucca*. Le due clausole sono state prodotte come aggiunte a un enunciato che avrebbe potuto terminare con *tutta la roba da mangiare e tutto*, il loro iniziare con il complementatore generico *che* mostra comunque l'intenzione di dipendere da una struttura sovraordinata. In questo modo Elisa ottiene una certa impressione di complessità:

- (47) ELI: || | { cioè } si vedono in questa casa.hhh  
 INV: mhm  
 ELI: con: tutta: la roba da mangiare e tutto # | che stanno bene | #che fhf c'è la mucca | che fa il latte || fhfhf [he

Diversamente, Valentina, la studentessa italiana che realizza il racconto più lungo e complesso, fa un uso ridotto di unità inferiori alla clausola (3,97%), a favore di clausole indipendenti (54,76%) e clausole subordinate (30,43%), queste ultime di vario tipo, argomentali, avverbiali e relative. Prevalgono le subordinate argomentali ed avverbiali, ridotto l'uso di relative.

A t2, emergono nelle due studentesse italiane alcuni cambiamenti, con una riduzione delle differenze individuali. Valentina ottiene valori vistosamente più bassi nella misura relativa alla subordinazione, mentre Elisa aumenta la complessità delle produzioni sia rispetto a lunghezza che subordinazione. La riduzione ottenuta da Valentina è dovuta ad alcune difficoltà della studentessa nel seguire il filmato video. Di conseguenza il suo racconto è poco preciso, poco dettagliato e soprattutto influenzato da interventi costanti dell'intervistatore che cerca di aiutarla a concludere il task. Questo dato mostra come la variazione possa essere legata in alcuni casi ad elementi esterni al task. A t2 Elisa dimostra invece diversi cambiamenti dal punto di vista sia quantitativo che qualitativo: pur mantenendo una certa essenzialità il suo racconto è un po' più lungo e dettagliato, e raggiunge valori di complessità maggiori che a t1. La studentessa produce clausole più lunghe e AS-unit più complesse, come illustrato dall'esempio (48). I cambiamenti quantitativi sono accompagnati da variazioni qualitative della distribuzione dei tipi di subordinate: si riduce l'uso di relative, a favore di avverbiali e argomentali.

- (48) ELI: || e:: # solo che alla fine:: si:: si vede | che:: e::hm praticamente il  
cacciatore e la pantera rosa si ritrovano insieme ancora | a scappare dal  
cane | che c'ha {la::} # la dinamite in bocca fhfhf ||  
INV: fhfhf fokeyf  
ELI: || {e poi:: e:: # e::hm ah} sì || #  
ELI: || poi dopo:: alla fine:: ehm {il cacciatore} # il cane riesce | a  
raggiungerlo || #  
ELI: || solo che:: gli mette la dinamite davanti ||  
ELI: || lui cerca | di coprirlo con della terra? || #  
ELI: || e:: solo che esplode tutto ||  
ELI: || e viene un buco nell'isola? || #0\_6  
ELI: || e {sprofonda::} fhfhf sprofonda nel mare ||  
ELI: || e si vede:: l'ultima scena | che si ritrovano su un pezzo di legno | a  
galleggiare:: ||  
INV: tutti insieme  
ELI: || e:: la pantera rosa è sempre dentro l'acqua con un ombrello ||  
INV: okey  
ELI: || che non ho capito perché. ||  
ELI: || e poi ci sono:: # il cacciatore e il cane su un pezzo di legno ||

Il task intervista è per sua natura più interattivo di quelli fin qui considerati. In questo caso intervistatore e apprendente sono coinvolti in una conversazione a due. In genere l'intervistatore guida l'interazione introducendo domande e lasciando alla studentessa lunghi turni di risposta, così come in (49) e in (50), dove si riportano due estratti prodotti rispettivamente da Elisa e Valentina a t1.

- (49) INV: mh. #0\_6 e che tipo di persone sono. se dovessi descrivermele?  
ELI: || allora:, #0\_4 mio padre è una persona un po':: in molto introverso,  
f\*sempre\*f zitto, || #  
INV: mh mh.  
ELI: || tranquillo. ||  
ELI: || non mi:: f\*cioè\*f insomma andiam d'accor. ||  
ELI: || mia madre invece è tutto il contrario ||  
ELI: || è, # sempre {di f estrov estrov} estroversissima. ||  
ELI: || parla con tutti, || #  
ELI: || son due persone oppostissime ||  
ELI: || ma:: van d'accordo. || #0\_6
- (50) INV: cosa gli fai fare?  
VAL: || di solito:/ guardiam la televisione insieme: || #  
VAL: || oppure a volte vuole | che gli faccio fare degli esercizi # tipo problemi:  
# dettati: così || #  
INV: !a::h! # come si chiama #  
VAL: || giuseppe || #  
INV: °giuseppe° # e:: quando guardate la televisione che cosa guardate insieme  
#  
VAL: || guardiamo: i fcartoni animati || #  
INV: i cartoni animati tu devi adattarti fallel  
VAL: || sì: || #



VAL: || oppure # e perché per esempio a volte mettono i canali della musica  
così || #

VAL: || solo che a lui non gli piace molto ||

VAL: || quindi: # guardiamo i cartoni animati. || #

In diversi punti però la conversazione porta a scambi più rapidi, è il caso in cui ad esempio l'intervistatore richiede approfondimenti, chiarimenti o commenta quanto detto dalla studentessa. Dal punto di vista sintattico il task favorisce l'uso di AS-unit semplici costituite di clausole indipendenti o unità indipendenti inferiori alla clausola e l'introduzione di un argomento per volta. La distribuzione qualitativa dei tipi di subordinate mostra come, nonostante l'uso quantitativamente ridotto, siano realizzati vari tipi di subordinazione.

In prospettiva longitudinale per Elisa non si registrano cambiamenti sostanziali rispetto alle misure di complessità, Valentina riporta invece una rapida riduzione della lunghezza media delle clausole. Questo è dovuto all'andamento del flusso conversazionale. Valentina a t2 risulta meno disposta a raccontarsi rispetto a t1, spesso risponde alle domande in modo vago, con una o due AS-unit semplici, richiedendo diversi interventi da parte dell'intervistatore e dunque maggiore interazione, turni più brevi, fatti di unità indipendenti inferiori alla clausola. Anche in questo caso le produzioni di Valentina mostrano come fattori esterni al task possono causare variazione individuale.

Infine, il task di apertura delle telefonate è quello che mostra maggior variazione situazionale rispetto alla complessità, ma minima variazione individuale: le due studentesse italiane realizzano in modo sostanzialmente simile il task. Vediamone alcuni esempi.

Le telefonate di servizio rappresentano un genere ben definito tra le conversazioni telefoniche, avendo lo scopo di richiedere e fornire dei servizi, quali informazioni, prenotazioni o assistenza. L'apertura è caratterizzata da una forte compressione delle mosse comunicative, che sono spesso ridotte all'essenziale e si basano su alcune formule di routine. La letteratura sulle aperture delle telefonate (ad es. Schegloff, 1979; 1986; per l'italiano Bercelli & Pallotti, 2002) ha evidenziato la presenza ricorrente di alcune mosse: (a)

apertura del canale: chi chiama manda un segnale (lo squillo del telefono) per verificare se c'è qualcuno dall'altro capo del filo; chi risponde, dicendo *pronto* o fornendo la propria identità, mostra che qualcuno c'è ed è pronto a interagire; (b) identificazione: il chiamante controlla se ha effettivamente raggiunto la persona desiderata ricevendo una risposta positiva; passa poi all'auto-identificazione, non sempre presente nelle chiamate di servizio a negozi, dove invece è spesso inclusa nella risposta del ricevente; (c) saluti: possono essere prodotti dal chiamante, dal ricevente o da entrambi; (d) interessanti: si trovano esclusivamente nelle telefonate tra intimi o in alcuni programmi radio-televisivi, mentre mancano nelle telefonate di servizio; (e) entrare nel merito: esauriti i cerimoniali di apertura, i partecipanti si orientano al motivo della telefonata, introducendo il primo topic della conversazione. La richiesta di informazioni è generalmente introdotta da pre-richieste e mitigata attraverso forme indirette o modificatori.

A t1 Elisa segue la struttura prima delineata: dopo l'apertura del canale, saluta, si auto-identifica ed entra nel merito mediante una serie di turni molto brevi, che introducono un argomento per volta e sono contingenti rispetto a quelli dell'interlocutore, creando così una punteggiatura piuttosto rapida e staccata del flusso conversazionale, come illustrato dall'esempio (51). Qui dopo le mosse iniziali, Elisa entra nel merito introducendo prima della richiesta vera e propria pre-richieste distribuite su diverse AS-unit e diversi turni. La densità informativa di ciascuna AS-unit è molto bassa e anche quando si introduce la richiesta vera e propria, essa è formulata nel modo più generico possibile.

- (51) ELI: ((squillo))  
 AG4: pronto?  
 ELI: || {e:} buongiorno. ||  
 AG4: sì #  
 ELI: || siamo dieci persone ||  
 ELI: || e dovremmo andare a fare: un viaggio a londra. || #  
 AG4: sì #  
 ELI: || volevo alcune informazioni? ||  
 AG4: sì.  
 ELI: || {hh allora e:: mh: come: cioè } in aereo {come::} # come:: siete organizzati || #0\_2

Valentina, realizza l'apertura delle telefonate di servizio in modo analogo, come illustrato nell'esempio (52). Dopo l'apertura del canale e i saluti, l'entrata in merito viene formulata in 4 turni e 4 AS-unit, di cui 2 costituite da unità indipendenti inferiori alla clausola, 1 da una clausola indipendente con verbo a costrutto complesso e 1 AS-unit che realizza una frase scissa con topicalizzazione. La richiesta è poi realizzata in modo molto generico.

- (52) VAL: ((squillo))  
 AG2: molto viaggi #  
 VAL: || | buonasera | ||  
 AG2: sì  
 VAL: || | volevo chiederle informazione | || #  
 AG2: sì  
 VAL: || | e: per Londra | || #  
 AG2: sì  
 VAL: || | una classe di: ragazzi | ||  
 AG2: mh mh  
 VAL: || | e: qualcosa di conveniente {che::} # che c'è | ||  
 AG2: allora intanto mi dovresti dire le date # il quantitativo di persone

L'analisi qualitativa della distribuzione del tipo di clausole (tabelle 18-19) conferma il quadro descritto finora. Le due studentesse utilizzano in prevalenza AS-unit brevi e semplici, fatte di clausole indipendenti e unità indipendenti inferiori alla clausola, fanno un uso limitato della subordinazione, e anche nel caso in cui una clausola subordinata viene prodotta, questa in genere non compare mai nei primi turni ed è sempre preceduta dalle mosse iniziali, apertura, identificazione e saluti, realizzate con turni brevi e incalzanti, che introducono gradualmente l'atto di richiesta.

In prospettiva longitudinale si conferma il quadro delineato fin ora. Si registra un lieve aumento della misura di subordinazione in tre telefonate del corpus. Queste si differenziano dalle altre per lunghezza ed elaborazione, così come esemplificato in (53) prodotto da Elisa a t2.

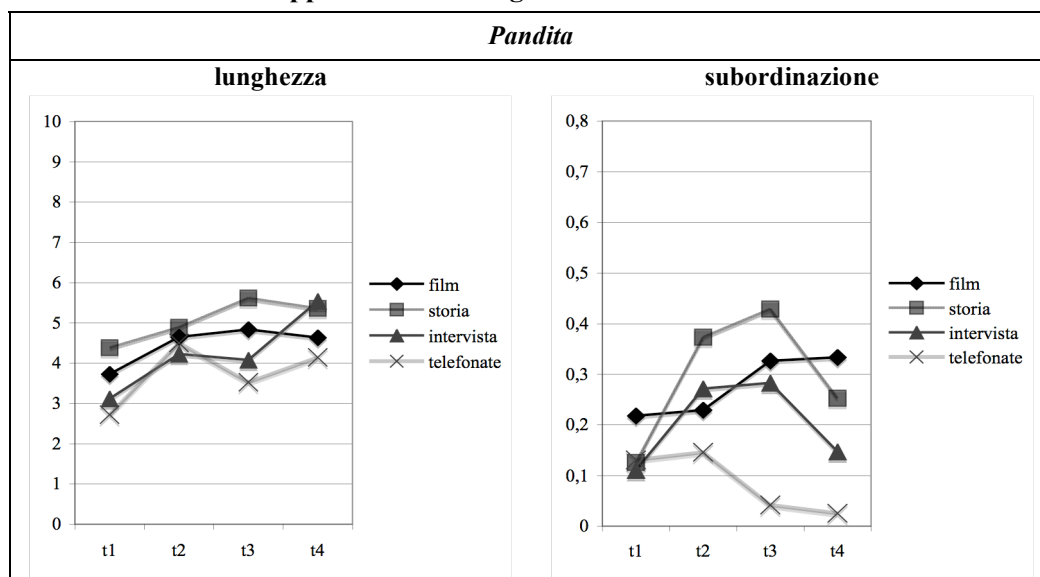
- (53) ELI: ((squillo))  
 SH1: mediastore buongiorno sono stefano  
 ELI: || | buongiorno | ||  
 ELI: || | volevo chiederle un'informazione | ||  
 SH1: mi dica  
 ELI: || | dovrei::: {consigliare} cioè regalare un cd ad una signora di quarant'anni | ||  
 SH1: sì

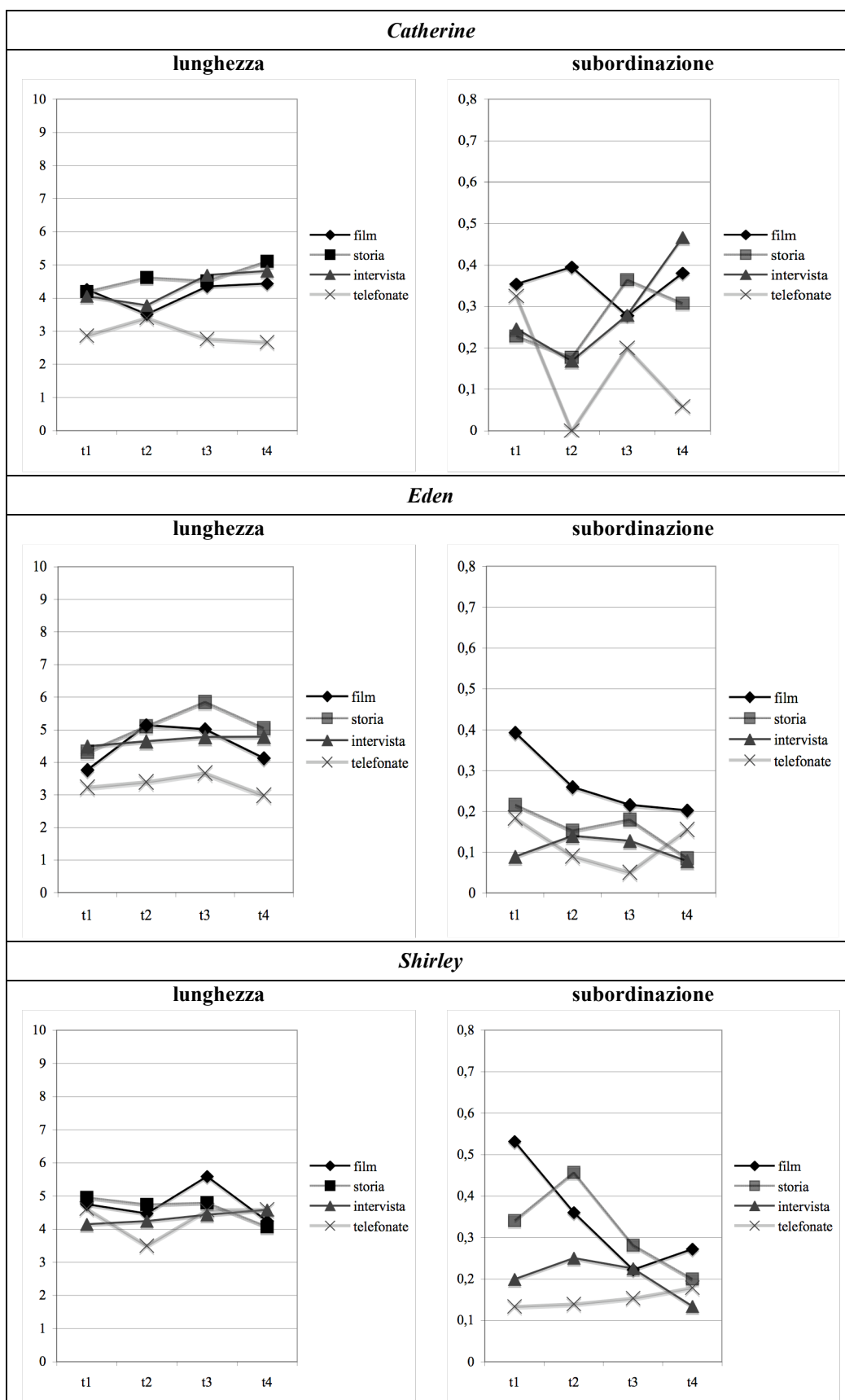
ELI: || | tch che::: comunque ama la musica # | ||  
 ELI: || | ascolta musica # da viaggio soprattutto | ||  
 ELI: || | e io avrei una lista di di di gruppi | ||  
 ELI: || | però non saprei::: | ||  
 SH1: ma una lista lunga  
 ELI: || | no::: | ||  
 ELI: || | son quattro gruppi | ||  
 ELI: || | però non so |{cosa:::} cosa sia meglio | || #  
 SH1: beh bisogna vedere che cosa ascolta lei generalmente  
 ELI: ma::: musica::: lei viaggia molto quindi:::

Dopo l'apertura dei canale e i saluti, Elisa entra nel merito e prepara la realizzazione della richiesta in 4 turni e 9 AS-unit. La richiesta è realizzata in modo generico e indiretto *però non so |{cosa:::} cosa sia meglio*, con una AS-unit di due clausole, di cui una costituita da una subordinata argomentale. Interessante osservare come anche in questo caso l'AS-unit complessa compaia solo al quarto turno.

Vediamo ora la variazione disegnata dalle apprendenti. Il grafico 8 e le tabelle 20-23 riportano i valori ottenuti dalle apprendenti nelle misure considerate. Una prima osservazione globale dei grafici ci permette da un lato di confermare il quadro di variazione identificato per il gruppo, dall'altro di osservare variazione individuale anche tra le apprendenti, soprattutto rispetto alla misura di subordinazione e a quelle relative alla distribuzione di tipi di clausole e di subordinate.

**Grafico 8. CAF – Complessità – Variazione individuale – Distribuzione per task – Apprendenti – Lunghezza e subordinazione**





**Tabella 20. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Pandita — Tipi di clausole**

	<i>t1</i>		<i>t2</i>		<i>t3</i>		<i>t4</i>	
<i>film</i>								
clausole indep.	127	66,49%	102	76,12%	134	68,72%	154	69,68%
unità inf. clausola	30	15,71%	7	5,22%	13	6,67%	11	4,98%
clausole sub.	34	17,80%	25	18,66%	48	24,62%	56	25,34%
avverbiali/sub.	6	17,65%	10	40,00%	25	52,08%	25	44,64%
argomentali/sub.	12	35,29%	11	44,00%	10	20,83%	11	19,64%
relative/sub.	16	47,06%	4	16,00%	13	27,08%	20	35,71%
<i>storia</i>								
clausole indep.	122	75,78%	58	71,60%	49	70,00%	82	66,13%
unità inf. clausola	21	13,04%	1	1,23%	-	-	17	13,71%
clausole sub.	18	11,18%	22	27,16%	21	30,00%	25	20,16%
avverbiali/sub.	5	27,78%	11	50,00%	13	61,90%	9	36,00%
argomentali/sub.	4	22,22%	6	27,27%	4	19,05%	3	12,00%
relative/sub.	9	50,00%	5	22,73%	4	19,05%	13	52,00%
<i>intervista</i>								
clausole indep.	80	49,38%	126	57,53%	179	52,19%	34	87,18%
unità inf. clausola	66	40,74%	46	21,00%	87	25,36%	-	-
clausole sub.	16	9,88%	47	21,46%	77	22,45%	5	12,82%
avverbiali/sub.	6	37,50%	15	31,91%	43	55,84%	5	100%
argomentali/sub.	3	18,75%	21	44,68%	16	20,78%	-	-
relative/sub.	7	43,75%	11	23,40%	18	23,38%	-	-
<i>telefonate</i>								
clausole indep.	2,17	48,22%	5,25	61,63%	3,43	52,15%	5,80	74,22%
unità inf. clausola	1,83	40,67%	2,50	27,12%	3,00	44,22%	1,60	23,56%
clausole sub.	0,50	11,11%	0,75	11,25%	0,29	3,63%	0,20	2,22%
avverbiali/sub.	x		-			x		-
argomentali/sub.	x		x			x		-
relative/sub.	-		x			-		x

**Tabella 21. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Catherine – Tipi di clausole e di subordinate**

	<i>t1</i>		<i>t2</i>		<i>t3</i>		<i>t4</i>	
	<i>film</i>							
clausole indep.	44	67,69%	32	60,38%	15	65,22%	118	64,13%
unità inf. clausola	4	6,15%	6	11,32%	3	13,04%	15	8,15%
clausole sub.	17	26,15%	15	28,30%	5	21,74%	51	27,72%
avverbiali/sub.	3	17,65%	3	20,00%	-	-	25	49,02%
argomentali/sub.	9	52,94%	8	53,33%	-	-	14	27,45%
relative/sub.	5	29,41%	4	26,67%	5	100,00%	12	23,53%
	<i>storia</i>							
clausole indep.	105	67,31%	97	72,93%	97	55,43%	167	70,17%
unità inf. clausola	22	14,10%	16	12,03%	31	17,71%	15	6,30%
clausole sub.	29	18,59%	20	15,04%	47	26,86%	56	23,53%
avverbiali/sub.	8	27,59%	7	35,00%	13	27,66%	33	58,93%
argomentali/sub.	16	55,17%	9	45,00%	29	61,70%	14	25,00%
relative/sub.	5	17,24%	4	20,00%	5	10,64%	9	16,07%
	<i>intervista</i>							
clausole indep.	90	54,22%	58	46,40%	86	65,15%	38	57,58%
unità inf. clausola	43	25,90%	49	39,20%	17	12,88%	7	10,61%
clausole sub.	33	19,88%	18	14,40%	29	21,97%	21	31,82%
avverbiali/sub.	19	57,58%	9	50,00%	14	48,28%	16	76,19%
argomentali/sub.	6	18,18%	6	33,33%	11	37,93%	2	9,52%
relative/sub.	8	24,24%	3	16,67%	4	13,79%	3	14,29%
	<i>telefonate</i>							
clausole indep.	1,75	29,58%	4,00	80,00%	3,17	38,67%	3,60	46,03%
unità inf. clausola	2,50	49,31%	1,00	20,00%	3,50	44,43%	3,60	48,89%
clausole sub.	1,50	21,11%	-	-	1,50	16,90%	0,40	5,08%
avverbiali/sub.	x		-		-		-	
argomentali/sub.	x		-		x		-	
relative/sub.	x		-		x		x	

**Tabella 22. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Eden – Tipi di clausole e di subordinate**

	<i>t1</i>		<i>t2</i>		<i>t3</i>		<i>t4</i>	
<i>film</i>								
clausole indep.	66	59,46%	75	74,26%	59	74,68%	84	69,42%
unità inf. clausola	14	12,61%	5	4,95%	6	7,59%	17	14,05%
clausole sub.	31	27,93%	21	20,79%	14	17,72%	20	16,53%
avverbiali/sub.	18	58,06%	13	61,90%	7	50,00%	6	30,00%
argomentali/sub.	8	25,81%	8	38,10%	4	28,57%	7	35,00%
relative/sub.	5	16,13%	-	-	3	21,43%	7	35,00%
<i>storia</i>								
clausole indep.	80	64,52%	94	77,69%	36	78,26%	51	79,69%
unità inf. clausola	22	17,74%	11	9,09%	3	6,52%	8	12,50%
clausole sub.	22	17,74%	16	13,22%	7	15,22%	5	7,81%
avverbiali/sub.	5	22,73%	5	31,25%	2	28,57%	2	40,00%
argomentali/sub.	7	31,82%	4	25,00%	2	28,57%	1	20,00%
relative/sub.	10	45,45%	7	43,75%	3	42,86%	2	40,00%
<i>intervista</i>								
clausole indep.	50	51,02%	115	55,02%	87	55,06%	38	55,07%
unità inf. clausola	40	40,82%	68	32,54%	53	33,54%	26	37,68%
clausole sub.	8	8,16%	26	12,44%	18	11,39%	5	7,25%
avverbiali/sub.	8	100,00%	16	61,54%	13	72,22%	4	80,00%
argomentali/sub.	-	-	4	15,38%	1	5,56%	1	20,00%
relative/sub.	-	-	6	23,08%	4	22,22%	0	0,00%
<i>telefonate</i>								
clausole indep.	3,00	52,67%	3,80	52,00%	6,00	66,67%	4,20	40,40%
unità inf. clausola	2,00	35,33%	3,00	40,67%	2,50	29,17%	4,80	46,23%
clausole sub.	0,60	12,00%	0,40	7,33%	0,50	4,17%	1,40	13,37%
avverbiali/sub.	x		x		-		x	
argomentali/sub.	x		x		-		x	
relative/sub.	-		-		x		x	



**Tabella 23. CAF – Complessità – Variazione individuale – Distribuzione per task  
– Shirley – Tipi di clausole e di subordinate**

	<i>t1</i>		<i>t2</i>		<i>t3</i>		<i>t4</i>	
<i>film</i>								
clausole indep.	75	61,98%	18	52,94%	14	63,64%	90	62,07%
unità inf. clausola	4	3,31%	7	20,59%	4	18,18%	24	16,55%
clausole sub.	42	34,71%	9	26,47%	4	18,18%	31	21,38%
avverbiali/sub.	27	64,29%	5	55,56%	3	75,00%	11	35,48%
argomentali/sub.	2	4,76%	2	22,22%	1	25,00%	9	29,03%
relative/sub.	13	30,95%	2	22,22%	-	-	11	35,48%
<i>storia</i>								
clausole indep.	41	65,08%	30	58,82%	52	63,41%	68	70,83%
unità inf. clausola	6	9,52%	5	9,80%	12	14,63%	12	12,50%
clausole sub.	16	25,40%	16	31,37%	18	21,95%	16	16,67%
avverbiali/sub.	5	31,25%	4	25,00%	4	22,22%	6	37,50%
argomentali/sub.	2	12,50%	2	12,50%	5	27,78%	7	43,75%
relative/sub.	9	56,25%	10	62,50%	9	50,00%	3	18,75%
<i>intervista</i>								
clausole indep.	203	60,24%	104	50,73%	140	52,83%	54	58,06%
unità inf. clausola	78	23,15%	60	29,27%	76	28,68%	28	30,11%
clausole sub.	56	16,62%	41	20,00%	49	18,49%	11	11,83%
avverbiali/sub.	30	53,57%	21	51,22%	18	36,73%	7	63,64%
argomentali/sub.	20	35,71%	12	29,27%	19	38,78%	3	27,27%
relative/sub.	6	10,71%	8	19,51%	12	24,49%	1	9,09%
<i>telefonate</i>								
clausole indep.	3,20	59,17%	3,25	27,92%	4,40	51,99%	7,20	59,51%
unità inf. clausola	1,80	29,17%	5,75	60,61%	2,80	36,19%	3,00	25,62%
clausole sub.	0,60	11,67%	1,25	11,47%	1,00	11,82%	1,80	14,87%
avverbiali/sub.	-		x		x		x	
argomentali/sub.	x		-		x		x	
relative/sub.	x		-		-		-	

A t1 Pandita e Catherine dimostrano poca variabilità nella complessità sintattica tra i quattro task. Catherine in particolare realizza una variazione con orientamento inverso rispetto alle italiane: il maggior valore nella misura di subordinazione è ottenuto nelle telefonate. Shirley e Eden al contrario dimostrano una maggiore variabilità tra task e già a t1 ottengono valori molto più vicini a quelli delle compagne italiane. Le differenze più notevoli riguardano anche in questo caso il task interattivo di apertura delle telefonate di servizio, dove le apprendenti utilizzano AS-unit più complesse rispetto alle italiane. In prospettiva longitudinale le traiettorie di sviluppo delle due coppie di apprendenti sono piuttosto diverse.

Pandita e Catherine, le meno avanzate a t1, dimostrano progressi significativi: c'è più variabilità tra task e la direzione di tale variazione è in sintonia con le produzioni delle studentesse italiane. Tale sviluppo è ottenuto attraverso una crescita della misura di subordinazione e una più varia distribuzione di tipi di subordinate nei task monologici, e una riduzione sia quantitativa che qualitativa nel task interattivo di telefonate. L'avvicinamento ai valori delle studentesse italiane avviene gradualmente, sebbene non sempre segua traiettorie lineari. Nei task monologici la complessificazione in una prima fase riguarda la clausola, con la produzione di un maggior numero medio di parole, in una seconda fase aumenta l'incidenza delle subordinate e la distribuzione per tipi è più simile alle modalità native. Nei task interattivi invece si assiste prima all'introduzione di più turni e poi alla distribuzione delle unità informative su diverse AS-unit semplici, con una conseguente riduzione della misura di subordinazione e della varietà dei tipi di subordinate impiegate.

Le traiettorie di sviluppo di Eden e Shirley, relativamente più avanzate già a t1, hanno un andamento diverso: per entrambe si registra nel tempo una graduale riduzione della misura di subordinazione in tutti i task, con una conseguente riduzione della variazione tra task. Nel caso di Eden questo è controbilanciato nei task monologici da un uso più vario e simile a quello nativo dei diversi tipi di subordinate, oltre che da una complessificazione a

livello di clausola, ottenuta grazie ad un aumento della misura di lunghezza, per Shirley invece è legato a scelte individuali rispetto alla strategia narrativa, basata più sull'intonazione calzante ed espressiva che sulla complessità. Per entrambe la riduzione per i task interattivi rileva invece diversi progressi a livello interazionale, in particolare in quello di apertura delle telefonate. Tra t1 e t4 Eden e Shirley realizzano scambi comunicativi man mano più rapidi e contingenti, diluendo la mossa di richiesta su più turni e distribuendo le informazioni su diverse AS-unit. In questo caso a t4 le apprendenti risultano un po' più complesse delle studentesse italiane, questo però non è da interpretarsi come un regresso, ma così come per le italiane a t2, l'analisi qualitativa mostra come esse utilizzino in realtà una maggior varietà di formule di richiesta o pre-richiesta contenenti in alcuni casi esempi di subordinazione.

Ricapitolando, i dati appena discussi confermano come la complessità sintattica possa aumentare in modo non lineare come conseguenza dell'interazione con livello, con task, tra misure e infine in relazione alla variazione individuale. Vediamo ora più in dettaglio il percorso di ciascuna apprendente attraverso alcuni esempi dal corpus. I paragrafi che seguono riportano, apprendente per apprendente, esempi delle produzioni realizzate nei diversi task. Si presentano prima i percorsi di Pandita e Catherine, poi quelli di Eden e Shirley.

A t1 la produzione di Pandita si discosta sotto diversi aspetti da quelle delle compagne italiane, con differenze sostanziali tra task monologici e interattivi.

Nei task monologici, Pandita realizza una narrazione molto lunga, ma caratterizzata da un andamento frammentario, per dettagli giustapposti. L'esempio (54) prodotto a t1 nel task film illustra come Pandita ricostruisce la stessa sequenza di eventi presa in considerazione per le compagne italiane.

- (54) PAN: || | e loro danno lavoro. | ||  
 INV: mh mh  
 PAN: || | { .h } c'è un.: # uomo | che dice | di prendere una: # legna {n} come  
       {co:} co:no #0\_6 | ||  
 INV: mh mh  
 PAN: || | fatta come un cono | ||

INV: .h come un cuneo  
 PAN: || si || #0\_5  
 PAN: || e li {pre-} da {funf fhf}  
 INV: .h  
 PAN: °un {bisto°} #1\_0 {°pis no. pistone}  
 PAN: || no.° || #0\_5  
 PAN: || {h} c'è {una-} # hh  
 PAN: || n[on so ||  
 PAN: || u una legna #0\_4 lunga ||  
 INV: [dimmi cosa vuoi dire # 0\_6 ah ah # un bastone  
 PAN: un bastone ||  
 INV: .h oka[y  
 PAN: || [e >ho detto< ||  
 PAN: || no. ||  
 PAN: || {lu} # l'uomo {dic} dice # ||  
 PAN: || no. ||  
 PAN: || non ho detto questo. || #  
 PAN: || ho detto come un ↑cono? ||  
 INV: ah ah  
 PAN: || {eh.} # in va | {dove ci s} dove {c'era} # fanno il nave ||  
 INV: mh mh  
 PAN: || c'era u:n bastone | #0\_3 che {c'è} >sotto< c'era una .h cono | #0\_5 che  
 ha fermato nave? ||  
 INV: fh[hf  
 PAN: || [fa prende da sotto?f ||  
 INV: fhf  
 PAN: || {fhf f} la nave lo partef ||  
 INV: fno f ho fh fhf .H  
 PAN: || [(va via) ||  
 PAN: || >ha detto< ||  
 PAN: || {o-} vado in {>piro<} prigionie fadesso f ||  
 INV: .h  
 PAN: || l[ui fpensa f ||  
 INV: [fhfhf ıh  
 PAN: || {poi} #1\_0 poi c'era una ragazza? | # che aveva fa:me | ||#0\_7  
 INV: e quindi cosa fa, se ne va?  
 PAN: || si. ||  
 PAN: || se ne va.  
 INV: a  
 PAN: da là. # cioè. || #0\_7  
 PAN: || poi c'era un'altra {u} una ragazza | che aveva fame #1\_0 ||  
 PAN: || loro due contrano ||  
 INV: e dove si trovano?  
 PAN: || in una: città ||  
 INV: ah ah  
 PAN: || { in: città #0\_3 vicino a un negosio #0\_2 del pane} no-. ||  
 PAN: || la ragazza | che aveva {fa fame} | .h ruba un pane # da panetteria cioè  
 fhf ||  
 INV: fh[hf  
 PAN: || [e scappa ||

La sequenza è lunga il doppio di quella delle compagne italiane, ma con un uso ridotto della subordinazione e AS-unit più brevi. A questa diversa distribuzione quantitativa corrisponde anche una diversa distribuzione qualitativa dei tipi di

subordinate: diversamente dalle compagne italiane Pandita utilizza prevalentemente relative (47,06%) o argomentali (35,29%) e in modo assai più limitato le subordinate avverbiali (17,65%). La maggior parte delle relative impiegate è introdotta dal presentativo *c'è*. L'effetto di frammentarietà è poi amplificato dalla quasi totale assenza di segnali discorsivi.

Tra t1 e t4 Pandita progredisce dimostrando un aumento complessivo delle misure di complessità, sia rispetto alla lunghezza che alla subordinazione. Tra t1 e t3 il progresso riguarda sia la lunghezza, con un incremento significativo del numero medio di parole per clausola, sia la misura di subordinazione. Tra t3 e t4 non ci sono incrementi in lunghezza e la misura di subordinazione decresce nel task storia. Le misure relative alla distribuzione del tipo di clausole e di subordinate mostrano come il cambiamento non sia solo di tipo quantitativo, ma soprattutto qualitativo, con una riduzione dell'uso di relative e una ridistribuzione delle clausole subordinate tra avverbiali e argomentali, con un avvicinamento al comportamento nativo. A t4 pertanto la narrazione è ricca di dettagli, più fluida e con un andamento meno frammentario, l'apprendente produce clausole più lunghe e AS-unit più complesse. Tale complessificazione quantitativa è supportata dall'uso di diversi tipi di subordinazione. Se a t1 la forma più rilevante di complessificazione sintattica era rappresentata dall'uso del *c'è* presentativo più relativa, a t4 Pandita utilizza una maggior varietà di tipi di subordinate, avvicinandosi alle scelte linguistiche delle compagne italiane, così come esemplificato in (55) realizzato a t4 nel task film.

- (55) PAN: || quindi il poliziotto ha scritto una lettera | # che è una persona (più)  
molto responsabile molto #0\_8 fidabile | || #  
PAN: || quindi:: # così trovi lavoro | ||  
INV: mh mh #0\_2  
PAN: || e poi prende quella lettera | ||  
PAN: || e va cercare un lavoro | || #1\_5  
PAN: || e va a cercare un lavoro {dove fanno} # praticamente | dove fanno  
#0\_5 le {barche::::} le navi::: | ||  
INV: si. #0\_2  
PAN: || .hh e lo trovano | || #  
PAN: || e lo danno lavoro | || #0\_2  
INV: mh mh #1\_5  
PAN: || il quale ha#nno detto di iniziare subito | ||

PAN: || {che lui} che lui però non voleva uscire da prigione | || #  
 PAN: || che lavoro devo farle? | || #  
 PAN: || dice | ||  
 PAN: || devi cercare {un} #0\_5 un legno come un triangolo | ||  
 INV: mh mh #0\_2  
 PAN: || e lui comincia a cercare | ||  
 PAN: || ma non lo trova | ||  
 INV: mh::: #0\_2  
 PAN: || no | || #  
 PAN: || vai a cercare da quella parte | || #  
 PAN: || da un'altra parte va cercare | || #  
 PAN: || lo trova {quel lì} # quel legno | che è fatto di triangolo però #0\_5 | che  
 viene usato per sostegno della nave | ||  
 INV: ah #0\_5  
 PAN: || se lo toglieva | andava via la nave | ||  
 PAN: || e andava nell'acqua! | ||  
 INV: ah ah #0\_2  
 PAN: || è quello | ||  
 INV: ah ah #0\_2  
 PAN: || ha cercato di toglierlo | ||  
 PAN: || e e ci riesce # a tirarlo | ||  
 INV: ah ah #0\_2  
 PAN: || e poi non guarda dietro | ||  
 PAN: || e porta quello | ||  
 INV: ah ah #0\_2  
 PAN: || però quando lo guarda | la nave va giù | ||  
 INV: ah ah #0\_2  
 PAN: || e::: #0\_6 e va nell'acqua # sotto l'acqua | ||  
 PAN: || e tutti guardano così | ||  
 INV: haha #0\_2  
 PAN: || {e lui} # lui scappa da là prima | che lo mettevano di nuovo in prigione  
 | || #  
 PAN: || perché lui voleva andare di nuovo in prigione | || #  
 PAN: || però #0\_5 scappa da là | ||

Nel task intervista a t1 la produzione di Pandita si discosta sia qualitativamente che quantitativamente dalla produzione delle studentesse italiane, in particolare rispetto alla misura di lunghezza. La studentessa riporta nei dettagli le sue esperienze, parlando sicuramente più delle compagne italiane, l'effetto generale della sua produzione è però, come per i task monologici, di estrema frammentarietà. Le AS-unit sono brevi e semplici, giustapposte tra loro, senza nessuna esplicitazione linguistica dei legami tra le diverse unità informative, così come illustrato dall'esempio (56).

- (56) INV: e cosa avete fatto il primo giorno di scuola. raccontami il primo giorno di scuola.# in italia.  
 PAN: || {mh.} #0\_7 {da} inizio da casa. hh | ||  
 INV: s[i.]  
 PAN: || [de]vo prendere auto. | || #

PAN: || e la non conoscevo. | quale devo prendere. | || #  
 PAN: || f\*ho detto\*f ? | ||  
 PAN: || bomporto? | ||  
 PAN: || f\*ho detto\*f | ||  
 PAN: || sì. | ||  
 PAN: || allora son salita | ||  
 PAN: || .h f\*son venuta\*f qua. | #  
 PAN: | perché ho fatto questa scuola. | ||  
 INV: sì. #  
 PAN: || e son entrata. | || #  
 PAN: || io avevo solo foglio scritto. #0\_6 terza effe. | || #  
 PAN: || f\*ho fatto vedere.\*f | || #  
 PAN: || dov'è fterza?[effe.f] | ||  
 INV: [fhhf]  
 PAN: || [mi hanno detto | ||  
 PAN: || su. su. su. | ||  
 PAN: || sono andata avanti. | ||  
 PAN: || son10 salita. | || #0\_2

Significativi i progressi in prospettiva longitudinale. La misura di lunghezza registra un aumento lineare della lunghezza media delle clausole: a t4 le unità sono lunghe il doppio rispetto a t1. Meno lineare lo sviluppo della traiettoria registrata dalla misura di subordinazione: tra t1 e t3 si registra un aumento importante dell'uso di subordinate, qui Pandita registra valori superiori addirittura a quelli delle compagne, a t4 si osserva però una riduzione, con un allineamento sui valori ottenuti dalle italiane. A t4 se da un lato continua la complessificazione sintattica a livello di clausola, si riduce l'incidenza di subordinate. Come esemplificato in (57), tale riduzione è compensata dall'uso di strutture pseudo-coordinanti introdotte da *che* o *perché* che le consentono di dare un effetto di dipendenza semantica tra gli enunciati. La conversazione pertanto risulta coesa e fluente, senza quell'effetto di frammentarietà che caratterizzava t1. Tale riduzione in complessità la avvicina quantitativamente e qualitativamente alle scelte delle italiane.

- (57) PAN: || studio secondo me fino adesso abbiamo...cioè | ||  
 PAN: || perché stiamo facendo dei ripassi per me | ||  
 PAN: || perché ho fatto chimica da 5 anni | ||  
 PAN: || sono le stesse cose | ||  
 PAN: || stiamo facendo dei ripassi | ||  
 PAN: || magari solo la matematica e fisica | ||  
 PAN: || che in superiore abbiamo fatto in primi 2 anni di fisica | non sapevo la lingua | ||  
 PAN: || quindi non # cioè # non ho studiato tanto bene | ||

PAN: || | poi ai {superiori devo stud/università} devo studiare quella | ||  
 PAN: || | ho studiato, | ||  
 PAN: || | ho già passata quella | ||  
 PAN: || | poi è andata tanto bene rispetto a altre materie | ||  
 PAN: || | sono andate più bene della chimica. | ||  
 INV: ah! allora sei proprio brava!

Infine, anche rispetto alle telefonate, a t1 Pandita si differenzia notevolmente dalle compagne. Se sono presenti le mosse fondamentali che caratterizzano l'apertura delle telefonate, i saluti, l'entrata in merito, a volte l'autoidentificazione mediante appartenenza categoriale, l'entrata in merito si realizza però in un unico turno, composto spesso da una sola AS-unit lunga e complessa. Le pre-richieste che le studentesse native realizzano con una serie di turni o frasi indipendenti, vengono più spesso incorporate in un'unica costruzione insieme alla formulazione della richiesta vera e propria, così come esemplificato in (58). Qui Pandita realizza l'entrata in merito formulando la richiesta in un'unica AS-unit di due clausole piuttosto dense da un punto di vista informativo: introduce subito un argomento specifico, *i costi dell'aereo*, dopo una pausa di #1\_3 completa l'unità sintattica e semantica aggiungendo *su andare a Londra*. La pausa in questo caso è co-costruita con l'interlocutore, i costi dell'aereo è infatti una clausola sintatticamente completa, ma non semanticamente, per questo l'operatore rimane in attesa.

- (58) PAN: ((squillo))  
 AG2: vacanza for you. sono paola?  
 PAN: || | .hh buonasera. | || hh .hh m-  
 AG2: pronto?  
 PAN: || | buonasera. | || #0\_8  
 PAN: || | io vorrei sapere #0\_4 i costi dei aereo. |  
       #1\_3 su andare a londra. | || #1\_6  
 AG2: per partire quando?

In questa, così come nelle altre telefonate realizzate a t1, Pandita utilizza non solo più subordinate delle compagne italiane, ma anche una maggior varietà di tipi, in particolare avverbiali di tipo causale.

In prospettiva longitudinale assistiamo invece a diversi progressi. La traiettoria di sviluppo segue un andamento lineare, aumenta un po' la lunghezza media delle clausole e si riduce significativamente la misura di



subordinazione, con un graduale allineamento ai valori delle studentesse italiane. Nell'esempio (59), realizzato a t2, dopo l'apertura del canale e i saluti, Pandita entra nel merito, utilizza forme di pre-richiesta e realizza enunciati con una minor densità informativa rispetto a t1. A differenza delle parlanti native qui la richiesta vera e propria è ancora specifica e condensa due informazioni molto precise, numero di pagine e prezzo. Il progresso più evidente è dato però dalla distribuzione delle unità su più turni, con la richiesta che non compare più nei primi, ma viene bensì preparata da altre mosse. Cambia anche la distribuzione qualitativa dei tipi di subordinate impiegati, non più principalmente avverbiali come nelle rilevazioni precedenti, ma piuttosto argomentali o relative, con un avvicinamento anche rispetto a questo parametro alle scelte delle compagne italiane.

- (59) PAN: ((squillo))  
 SH1: pronto  
 PAN: || pronto ||  
 PAN: || buonasera ||  
 SH1: buonasera  
 PAN: || io volevo chiedere # un'informazione su un libro ||  
 SH1: dica  
 PAN: || è un libro di:: autore di Danielle Steel ||  
 SH1: sì::  
 PAN: || il titolo è il {renz}  
 SH1: °il ranch°  
 PAN: {ranch} # ranch ||  
 SH1: ah  
 PAN: || e # volevo sapere:: | quante pagine ci sono | e:: quanto costa ||  
 SH1: no un titolo del genere Danielle Steel non lo conosco

A t3 e t4, Pandita realizza la telefonata mantenendo uno stile sintatticamente semplice, con uno scambio molto più vicino a quello delle parlanti native: le informazioni sono introdotte una per volta attraverso AS-unit semplici, le mosse tipiche dell'apertura distribuite su più turni, così come esemplificato in (60) realizzato a t4.

- (60) PAN: ((squillo))  
 AG2: buonasera agenzia (dila)ventura::?  
 PAN: || buonas::era || #0\_8  
 PAN: || scusi ||  
 PAN: || vorrei sapere informazioni:: || #0\_5  
 PAN: || perché vorrei fare una vacanza a milano || #1\_5  
 AG2: no # non abbiamo mica niente # su una vacanza a milano #0\_5

PAN: || sì. ||  
 AG2: no. #0\_5  
 PAN: || | cioè non conosce degli alberghi::: || #0\_5  
 AG2: no. # no signora non conosciamo alberghi #0\_3  
 PAN: niente::?  
 AG2: no no no  
 SHI: °allora cosa fate°  
 AG2: facciamo l'estero no # milano  
 PAN: ah::: okey grazie  
 AG2: prego

Vediamo ora il percorso individuale di Catherine. Nei task monologici, a t1 Catherine produce un racconto che per lunghezza totale e valori di complessità si situa entro il quadro di variazione delineato dalle due studentesse italiane. Emergono però alcune differenze da un punto di vista qualitativo, così come illustrato dall'esempio (61) realizzato nel task film a t1. Catherine racconta il video con una certa imprecisione, fornendo piuttosto un elenco di eventi più o meno principali dove i fatti non sono sempre esplicitamente collegati tra loro. L'analisi della distribuzione dei tipi di clausole mostra come le subordinate utilizzate siano soprattutto argomentali e relative, piuttosto che avverbiali.

- (61) CAT: || | un giorno >è andato< in un negossio | ||  
 CAT: || | (ha) incontrato una: {una} donna #3\_0 | che {i signore/ # rubato u::n}  
 signore rubato: #0\_4 un panino  
 CAT: || | non lo so | ||  
 CAT: un baguette | || #0\_2  
 CAT: || | e hanno da- {hann} hanno dato la colpa alla signora | || #1\_5  
 INV: mhmm\ #3\_0  
 CAT: || | {e:::} #1\_6 e::: {quell che ve} quella | che vende il panino, | .hh ha  
 chiamato la: polisiotta | per venire a arrestarlo | || #1\_2

In prospettiva longitudinale dimostra anche lei diversi progressi, ma diversamente da Pandita la misura di lunghezza e di subordinazione aumentano con traiettorie non lineari a zig zag, con effetti di trade-off nelle due misure e andamenti opposti nel film e nella storia. A t4 Catherine ottiene i progressi più significativi. Le sue produzioni nei due task sono lunghe il doppio o il triplo delle precedenti, il racconto risulta più coerente, preciso e dettagliato e Catherine sembra svolgere con più agio il task. Quantitativamente aumentano entrambe le misure di complessità: subordinazione e lunghezza. Progressi anche più significativi si ottengono rispetto ad alcuni aspetti qualitativi della produzione. Qui la narrazione è più fluida, precisa e gli eventi vengono

ricostruiti in modo più chiaro che a t1. La distribuzione del tipo di subordinate è ora più vicina a quella delle studentesse italiane: aumenta in modo significativo l'uso di subordinate avverbiali, come esemplificato in (62), realizzato nel task film a t4.

- (62) CAT: || | {allora lui #2\_5 tch uscendo da lavoro:: oh} mentre lui {f\*andava andava} andava | {c'era una\*f signorina #0\_5} c'era una signorina #0\_5 .hh | che ha visto mh un:: panettiere | ||  
 INV: mh:::  
 CAT: || | si dice un panettiere | || #0\_2  
 CAT: || | [un panettiere | che tirava delle pane dal:: sua macchina | {per} per portare:: # mh:: {negossio} [[suo negozio | ||  
 INV: [sì sì  
 INV: [[negozio mh mh #  
 CAT: || | .hhh e la ragazza avendo fame:: ehm:: | quando ha visto | che uomo è entrato in negozio | poi aveva:: il furgone aperto | .hh lei l'ha preso il panino | ||  
 CAT: || | e ha iniziato a correre | ||  
 INV: mh mh  
 CAT: || | {essendo che lei fhhf ehm:: #0\_6 cioè} quando {lui} lei ha iniziato a correre | è entrato {una signori-} una signora | || #0\_5  
 CAT: || | {l'ha vist} l'ha vista | che:: la ragazza ha preso il pane | per scappare lì | ||  
 CAT: || | {l per:: mh::} #0\_5 ha preso il pane: | ||  
 CAT: || | e:: #0\_5 sta scappando | ||  
 INV: mh mh  
 CAT: || | .hh appe- allora appena è arrivato:: uomo # | panettiera l'ha detto | ||  
 CAT: || | guarda | che c'è ragazza lì | che sta correndo | ||  
 CAT: || | ha preso {panino} # .hhh sua baguette | ||  
 CAT: || | era baguette | || #0\_2  
 CAT: || | ha preso tua baguette | ||

Nel task intervista a t1 Catherine realizza questo task senza discostarsi in modo significativo dalle scelte delle compagne italiane, così come evidente da (63).

- (63) INV: cosa hai pensato quando hai saputo che saresti venuta in italia #0\_9 a stare con i tuoi #0\_6  
 CAT: || | cioè io ero contenta | perché:: #0\_4 finalmente adesso vivo con miei genitori #0\_8 | perché non sono mai {abitu} cioè abitata con loro | || #0\_8  
 CAT: || | sono sempre l- qua | || #0\_6  
 CAT: || | vengono in ghana | || #0\_3  
 CAT: || | fanno un mese | ||  
 CAT: || | e poi tornano qua | || #0\_6  
 CAT: || | allora #2\_1 ehmm #2\_0 era ehm #0\_3 ero un po' contenta #0\_5 di #0\_7 | ||  
 INV: un po' o molto #0\_3  
 CAT: || | °£molto anche°£ | || #1\_4

Il suo sviluppo longitudinale si differenzia invece da quello realizzato dalle

altre apprendenti. Se infatti la misura di lunghezza non dimostra variazioni sostanziali, quella di subordinazione cresce in modo evidente tra t2 e t4, dove raggiunge valori due volte superiori a quelli delle compagne italiane. A differenza delle compagne, Catherine complessifica gradualmente le sue produzioni, usando più subordinate e riducendo la quantità di unità indipendenti inferiori alla clausola, come esemplificato dai due estratti (64) e (65) realizzati rispettivamente a t3 e t4.

- (64) CAT: || allora ero molto contenta | di vivere coi miei | perché: là abitavo con la mia zia | || #  
 CAT: || e:: # ero molto contenta | quando venivo qua | || #  
 CAT: || almeno se non c'è nessuno | che conoscevo | {av-} i miei # e:h {co-conos-} conoscevano | ||
- (65) CAT: || ma il vantaggio dalla scuola all'università è | che puoi fare l'esame | quando ti pare | ||  
 CAT: || vai a lezione quando ti pare | ||  
 CAT: || però qua sei obbligato a fare l'esame | che ci sono | ||  
 CAT: || e poi se passi o non passi | voto ti mettono lì | ||  
 CAT: || e poi qua sei obbligato a venire a scuola sempre | ||  
 CAT: || se no se fai troppi assenti | ti chiamano a casa i tuoi: | ||  
 CAT: || gli dicono | ||  
 CAT: || però all'università niente | ||  
 CAT: || e poi qua non è che devi studiare dalla mattina fino alla sera: | ||  
 CAT: || ti basta 4 ore | per studiare | ||  
 CAT: || però all'università ci vuole molto, molto, molto, molto da studiare, molto | ||  
 CAT: || non devi neanche perdere un secondo | per fare ogni cosa | ||

Rispetto al task telefonate, a t1 invece le produzioni di Catherine differiscono notevolmente da quelle realizzate dalle compagne italiane, in particolare rispetto alla misura di subordinazione: così come Pandita, anche qui tutte le informazioni sono introdotte in un solo pacchetto che risulta così composto da un cospicuo numero di parole e clausole. In (66) possiamo vedere un esempio delle aperture realizzate da Catherine.

- (66) CAT: ((squillo))  
 SH3: gigatel buonasera sono [clau in che cosa  
 CAT: || [{e:} pronto | || #  
 CAT: || {buon-} buonasera | ||  
 SH3: buonasera # [mi dica  
 CAT: || [{vorrei- eh} vorrei avere informazioni | se avete ericson zeta seicento | || #  
 CAT: || il telefono di ericson | || #2\_2

Qui è interessante notare l'assenza di sincronia nei primissimi turni in cui si realizzano l'apertura del canale e i saluti: Catherine non lascia che l'addetta finisca la sua frase di routine *sono Claudia in che cosa posso esserle utile?*, ma entra a metà con il suo *pronto buonasera*, causando così un piccolo scompenso nel flusso conversazionale. Per entrare nel merito, in questa così come nelle altre telefonate a t1, Catherine usa la formula *vorrei avere informazioni se avete*, non solo scorretta da un punto di vista grammaticale, ma anche del tutto atipica sul piano pragmatico: abbiamo visto come le studentesse italiane di solito producono *vorrei avere un'informazione* come unità a sé, a cui segue una nuova frase indipendente, frequentemente sotto forma di domanda diretta. Catherine, così come le altre apprendenti, realizza la richiesta con una AS-unit complessa, composta di due clausole e sei parole (considerando una parola unica marca e modello del cellulare).

In prospettiva longitudinale Catherine ottiene diversi progressi. La sua traiettoria di sviluppo non ha un andamento lineare come quello della compagna Pandita, ma procede piuttosto con un'alternanza di progressi e di regressi. A t3 l'analisi quantitativa pare mostrare regresso in quanto assistiamo a un aumento delle misure di complessità, l'analisi qualitativa mostra però come tale regresso sia solo apparente. L'aumento dell'incidenza di clausole subordinate è controbilanciato da un avvicinamento alle modalità native nella scelta del tipo di clausola da utilizzare. Possiamo osservare nell'esempio (67) come Catherine realizzi l'apertura rispettando le modalità native, cioè producendo AS-unit brevi e contingenti con una bassa densità informativa. Le due AS-unit complesse realizzate a t3 sono infatti ben diverse da quelle impiegate a t1. Si tratta in questo caso di due relative che hanno la funzione di specificare l'entità dell'oggetto su cui verterà la richiesta. Abbiamo visto in precedenza come esempi di questo tipo siano presenti anche nel corpus delle italiane. Interessante notare come l'AS-unit complessa sia distribuita su due turni, mantenendo quindi una certa rapidità nello scambio comunicativo.

(67) CAT: ((squillo))  
 SH3: ricordi mediastore buongiorno sono marco  
 CAT: || pronto ||  
 CAT: || buongiorno || # .hhh  
 SH3: prego  
 CAT: || eh mi scusi ||  
 CAT: || mi devo chiedere un'un'informazione ||  
 SH3: prego  
 CAT: || l'ultimo cd di norah jones |  
 SH3: sì  
 CAT: | che si intitola not too [late ||  
 SH3: [too late # sì sì ce l'abbiamo  
 CAT: || quanto costerà secondo lei ||  
 SH3: e::: uf dovrebbe essere venti euro e novanta

A t4 le produzioni di Catherine sono sia qualitativamente che quantitativamente più vicina a quelle delle studentesse italiane. L'apertura è realizzata con AS-unit brevi, la richiesta distribuita su più turni, mitigata da forme di pre-richiesta e formulata in modo generico.

(68) CAT: ((squillo))  
 EA1: pronto? #0\_5  
 CAT: || pronto || #  
 CAT: || buongiorno || #1\_5  
 EA1: buongiorno  
 CAT: || è la [signo- ||  
 EA1: [chi parla?  
 CAT: || la catherine || #1\_8  
 EA1: ah ciao #0\_5 dimmi  
 CAT: || allora #0\_5 {volevo} # volevo chiederti un'informazione || #1\_0  
 EA1: sì  
 CAT: || su un viaggio a verona || #1\_0  
 EA1: sì. #0\_2 da dove devi partire::?

Vediamo ora le produzioni realizzate da Eden. Nei task interattivi, a t1 produce racconti già vicini al quadro di variazione identificato per le studentesse italiane, dimostrando dunque di essere già piuttosto avanzata fin dalla prima raccolta dati. Le sue produzioni sono lunghe e dettagliate, caratterizzate da un uso importante della subordinazione, soprattutto nel film, dove Eden riporta il racconto ricostruendo in modo coerente e lineare gli eventi più importanti. L'estratto in (69) prodotto nel task film a t1 riporta una delle sequenze centrali del video, qui Eden produce diverse AS-unit complesse, utilizzando subordinate di vario tipo, in particolari avverbiali e argomentali.

Anche la distribuzione qualitativa del tipo di clausole è vicina a quella delle italiane.

- (69) EDE: || mentre va in strada così | #0\_2 c'è {una:} #0\_2 {una signora}  
 EDE: || non so ||  
 EDE: {una:} #0\_3 (u)na ragazza, .hh | che: {va} #0\_8 entra in un negozio | ||  
 EDE: || prende il pane | ||  
 EDE: || inizia a pagare | ||  
 EDE: || però scappa. .h | ||  
 EDE: || mentre sta scappando | incontra charlie chaplin: #0\_2 sul marciapiede |  
 ||  
 EDE: || si incontrano | ||  
 EDE: || cadono insieme. .h | ||  
 EDE: || allora la signora va- a chiamare: .hh quel signore il proprietario, # | ||  
 EDE: || però il proprietario #0\_4 {cioè} prende lui | dicendo | che è stato lui # |  
 ||  
 EDE: || {f\*cioè} # insomma {la d-} \*f charlie chaplin la difende .h | dicendo |  
 che sono stato io | che {ho ruba:} ho rubato questo [pane qua] | ||  
 EDE: || dice. .hh | || #0\_2  
 EDE: || va beh alla fine? #0\_7 la signora | che ha visto: #0\_4 questa ragazza  
 #0\_3 | dice | ||  
 EDE: || è stata lei | ||  
 EDE: || è stata lei | ||  
 EDE: || non è stato lui. | || #0\_3  
 EDE: || va beh | ||  
 EDE: || alla fine prendono lei, | || #1\_0

In prospettiva longitudinale nel suo caso il progresso è realizzato attraverso una graduale riduzione della misura di subordinazione e un aumento della lunghezza media delle clausole, anche in questo caso con una traiettoria di sviluppo non lineare, con un effetto di trade-off tra le due misure. La riduzione quantitativa della misura di subordinazione è qui bilanciato da un uso qualitativamente più vario di tipi di subordinate che aiutano l'apprendente a strutturare con più precisione gli eventi, così come esemplificato in (70).

- (70) EDE: || poi dopo # mentre va nella strada così #1\_0 | incontra una ragazza | ||  
 #0\_8  
 EDE: || cioè no che incontra:: | || #2\_0  
 EDE: || nel frattempo fanno vedere una ragazza #1\_2 | che cammina per la  
 strada |  
 INV: sì:: #0\_2  
 EDE: che arriva {dal:: #0\_8} davanti al forno:: | che vede:: | || #2\_0  
 EDE: || insomma ha fame | ||  
 INV: mh mh #0\_5  
 EDE: || allora ruba questo pane | ||  
 EDE: || e:: comincia a scappare | ||  
 EDE: || allora una signora l'ha vista:: | || #0\_8  
 EDE: || va dal fornaio | ||  
 EDE: || e dice | che quella::: ragazza (niente) ha rubato il pane | ||

EDE: || allora #0\_5 si incontrano con il charlie chaplin | ||  
 INV: ah ah #0\_2  
 EDE: || è cadono per terra | ||  
 INV: la ragazza? #0\_2  
 EDE: || la ragazza | ||  
 EDE: || sì:: | ||  
 INV: mh mh  
 EDE: || charlie chaplin {rimane::} #0\_8 gli rimane il pane in mano | ||  
 INV: sì::  
 EDE: || allora arrivano i poliziotti | ||  
 EDE: || e £prendono lui£ | || hehehe  
 EDE: || allora la signora cerca {di con- ehm} di convincere il fornaio | dicendo  
 | che è lei | || #0\_8  
 EDE: || allora corrono tutti quanti | || #0\_5  
 EDE: || {dei poliziotti poi #0\_5} un poliziotto:: #0\_2 prende la ragazza | ||  
 EDE: || e lasciano andare lui | || #

Anche nel task intervista, a t1 la produzione di Eden in termini di complessità è equivalente, sia qualitativamente che quantitativamente, a quelle delle compagne italiane. Rispetto alle altre apprendenti però la sua produzione risulta più semplice nella misura di subordinazione. Eden in questo task preferisce utilizzare AS-unit semplici, ma di una certa lunghezza, diluendo così la densità informativa dello scambio comunicativo. In prospettiva longitudinale non ci sono cambiamenti sostanziali tra t1 e t4 rispetto alle misure considerate, così come evidente dagli estratti (71) e (72), realizzati rispettivamente a t1 e a t4.

- (71) INV: e com'è stato il primo giorno di scuola  
 EDE: || bruttissimo | || [hehehe]  
 INV: [£dai raccontamelo£]  
 EDE: || bruttissimo | perché: f\*c'era il prof di matema | ||  
 EDE: || avevamo la prima ora | ||  
 EDE: || avevamo matematica\*f cioè | || #  
 EDE: || tutta la giornata è finita {a me} per me e per mia sorella | per spiegare  
 tutte le cose insomma | ||  
 EDE: || come funziona cioè | || .h  
 EDE: || almeno per creare un po' di: #1\_0 di confidenza, | ||  
 INV: !certo!  
 EDE: || sì | ||  
 EDE: || è finita tutta l'ora | ||  
 EDE: || {a noi} cioè e noi siamo arrivate a casa che [hehehe] | || #  
 INV: [eravate distrutte hh] #  
 EDE: || sì sì | || #0\_2  
 EDE: || troppo. | ||
- (72) INV: ho capito # ah #2\_4 va bene # senti #2\_2 e::: dimmi un po' tu adesso  
 lavori quand'è che hai finito la scuola? #  
 EDE: || l'anno scorso | || #



INV: l'anno scorso hai appena finito ah quindi hai trovato subito il lavoro  
 EDE: || no no io:: # lavoravo || #0\_2  
 EDE: || e poi venivo a scuola serale ||  
 INV: ah facevi il serale quindi  
 EDE: || perché io sono andata:: a fare lo stage là  
 INV: ah ah  
 EDE: e:: a novembre del 2006 || #  
 INV: sì  
 EDE: || poi mi hanno chiesto di lavorare con loro? || #0\_2  
 EDE: || e sono rimasta || #0\_2  
 EDE: || {e faccio cioè lavoro} # lavoravo di giorno || #  
 EDE: || facevo otto ore e poi la sera || #0\_2  
 INV: che brava  
 EDE: || facevo il corso ||

Infine, rispetto alle telefonate, anche per Eden a t1 le produzioni in questo task differiscono da quelle delle compagne italiane: l'entrata in merito, con la pre-richiesta e la successiva richiesta, sono realizzate anche qui attraverso AS-unit complesse dal punto di vista sintattico e dense dal punto di vista informativo, con la richiesta che indirizza un'informazione molto precisa, come esemplificato in (73). L'uso di un maggior numero di subordinate è accompagnato dalla produzione di diversi tipi di subordinate, anche avverbiali (tabella 22).

- (73) EDE: ((squillo))  
 AG2: centro viaggi !  
 EDE: || buonasera ||  
 EDE: || volevo chiedere u(n)' informazione | per andare a barcellona .| ||  
 AG2: s::i ! un att(imo) , ma che tipo di informazioni lei voleva ?  
 EDE: || volevo: chiedere cio(è) | se c'è il tre::no e::h || +/.  
 AG2: →treno? no # non facciamo treno noi eh?

In prospettiva longitudinale anche Eden realizza diversi progressi. Tra t1 e t3 i valori delle misure di complessità raggiungono livelli identici a quelli delle parlanti native, grazie a un lieve aumento della lunghezza media delle clausole e un'importante riduzione quantitativa e qualitativa dell'uso di subordinate. Tra t3 e t4 sembra invece esservi un'inversione di tendenza delle due misure di complessità che ritornano verso i valori iniziali. Vediamo più in dettaglio come si sviluppano nel tempo le produzioni di Eden utilizzando alcuni esempi dal corpus.

A t2 Eden non ottiene cambiamenti vistosi. Questo è dovuto alle diverse modalità con cui realizza le richieste. In alcuni casi l'apertura della telefonata

non sembra essere molto diversa da t1. Dunque, a differenza delle studentesse italiane, non solo condensa più informazioni in un'unica AS-unit complessa contenente la pre-richiesta e la richiesta, ma come nel caso di Catherine, formalmente e pragmaticamente inusuale: *volevo chiedere::: informazione per un libro | se c'avete o no*. Abbiamo già precedentemente visto come in questi casi le studentesse italiane preferiscano produrre due AS-unit e soprattutto realizzare la richiesta con un'interrogativa diretta. In altri casi, Eden è in grado di condurre lo scambio di routine di apertura con enunciati rapidi e distaccati, dunque più vicini alle modalità delle studentesse italiane. A t3 non riscontriamo più tale variabilità nelle modalità di apertura e come esemplificato in (74) l'apertura viene realizzata con modalità simili a quelle delle native. Lo scambio comunicativo è rapido e contingente, la mossa di richiesta è distribuita su più turni e realizzata con AS-unit semplici. Anche la distribuzione qualitativa dei tipi di clausole è ora più in linea con i valori delle studentesse italiane.

- (74) EDE: ((squillo))  
 SH4: libreria giunti # buongiorno  
 EDE: || sì ||  
 EDE: || buongiorno ||  
 EDE: || sono eden || #  
 EDE: || vendete cd || #  
 SH4: e::hm # qualcuno  
 EDE: || qualcuno || #  
 EDE: || allora le dico i titoli::: || #  
 EDE: || non lo so || #  
 EDE: || elton john || #  
 SH4: e::: vediamo # xxxx un po'::: xxx guardo un attimo eh

A t4 le misure di complessità ritornano su valori vicini a quelli iniziali. L'aumento della misura di subordinazione è legato, in questo caso, all'uso di 4 AS-unit complesse. Un'osservazione qualitativa dei dati mostra però come il possibile regresso indicato da tale aumento sia solo apparente. L'aumento dell'incidenza di clausole subordinate è infatti legato all'uso di una maggior varietà di formule di richiesta nelle chiamate ai negozi, così come esemplificato in (75), e all'uso di AS-unit complesse nelle pre-richieste delle telefonate indirizzate agli esperti, così come esemplificato in (76) e (77). Così

come per le italiane le AS-unit complesse non compaiono più ai primissimi turni e sono anticipate da mosse preparatorie

- (75) EDE: ((squillo))  
 SH3: libreria (memoli) buonasera sono tatiana #0\_2  
 EDE: || eh:: buonasera || #0\_5  
 EDE: || allora # voi vendete divudi::? || #0\_5  
 SH3: sì::.  
 EDE: || sì:: || #  
 EDE: || okay || #  
 EDE: || volevo sapere | se avete:: # la finestra sul cortile || #8\_0  
 EDE: || °l'altro qual'era?° || #0\_8  
 ST4: °o l'assassinio sull'oriente express° #16\_0  
 SH3: sì. #0\_5 ce ne sono un po' di copie .....
- (76) EDE: || io l'ho chiamata | perché dovrei fare un regalo ad una signora {che ha}  
 # di quarant'anni ||
- (77) EDE: || allora io avrei bisogno di una mano:: mh # | da regalare un film giallo  
 a una signora || #0\_6

Vediamo infine il percorso di sviluppo seguito da Shirley. Nei task monologici, a t1 Shirley racconta la storia con un resoconto breve, poco denso di particolari, ma con valori di complessità piuttosto alti sia rispetto alla lunghezza che alla subordinazione. Nella sua produzione è importante l'uso di clausole subordinate: le informazioni tendono ad essere impacchettate in AS-unit lunghe e complesse. Qui i valori dell'analisi quantitativa tendono a raggiungere i livelli di Valentina, la studentessa italiana che parla di più.

Nell'esempio (78), prodotto a t1 Shirley racconta la sequenza del video considerata per le italiane:

- (78) SHI: || { .h # gli: hanno:: preso. # un lav- } a lavoro il primo giorno gli hanno  
 detto | di cercare un pezzo di: #0\_2 legno. ||  
 SHI: || .h stava cercando cercando cercando ||  
 SHI: || non l'ha trovato. ||  
 SHI: || ne ha trovati: #1\_0 una? ||  
 SHI: || che:: .h {e::} appena arrivato lì | fa, ||  
 SHI: {uh. #0\_5} || | caspita. ||  
 SHI: || guardalo qui. || #0\_6  
 SHI: || invece? # non era solo: #1\_0 un legno. ||  
 SHI: || ma serviva? | .h a tener ferma la:: b- barca | per non partir {da:: .h}  
 #0\_6 da solo. ||  
 INV: mhm  
 SHI: || appena: # l'ha tolto? | parte la: #0\_2 barca. ||

SHI: || e:: # da lì l'hanno licenziato. ||

Rispetto alla distribuzione qualitativa dei tipi di subordinate, nella sua produzione è significativamente superiore la presenza di avverbiali (64,29%) e relative (30,95%), a scapito delle argomentali (4,76%).

In prospettiva longitudinale, i resoconti del video prodotti a t2 e t3 risultano essere tre volte più brevi e meno dettagliati del racconto prodotto a t1. A t4 la narrazione torna ad essere un po' più precisa, ma mantiene sempre un maggiore grado di semplicità rispetto a t1. In modo più evidente a t2 e t3, Shirley si limita a riportare in modo molto sommario la trama, senza scendere negli dettagli delle azioni, cercando piuttosto di portare a termine il compito il prima possibile e avvicinandosi così alle modalità di narrazione tipiche di Elisa, la studentessa italiana che produceva il resoconto più semplice. La scelta stilistica di una narrazione snella ed essenziale si riflette nella progressiva riduzione delle misure di complessità, in particolare di quella relativa alla subordinazione. La misura di lunghezza non dimostra cambiamenti importanti, dopo una lieve riduzione a t2, un aumento a t3, a t4 Shirley ritorna agli stessi valori ottenuti a t1. Più evidenti invece i cambiamenti legati alla misura di subordinazione: tra t1 e t3 si assiste a una riduzione evidente dell'uso di subordinate, seguito poi da una nuova lieve crescita a t4. All'ultima rilevazione il racconto risulta essere realizzato favorendo l'uso di AS-unit semplici, la coerenza nella narrazione è ottenuta attraverso la giustapposizione di clausole indipendenti e frammenti sostenuti da un'intonazione calzante ed espressiva. I cambiamenti ottenuti da Shirley, già avanzata a t1, sono dunque legati non tanto ad uno sviluppo della complessità sintattica intesa come acquisizione di forme di subordinazione più complessa o unità di produzioni più lunghe, ma sembrano piuttosto dovuti alla sperimentazione di diversi stili discorsivi.

- (79) SHI: || e subito # gli:: #0\_8 prendono || #  
SHI: || e:: #0\_5 comincia subito || #0\_5  
SHI: || solo che gli:: # dicono | di andare a prendere un:: tipo di:: # {le} legno |  
||  
INV: mh mh  
SHI: || .hh che lui comunque ||  
SHI: || sì [fokey£ || #  
INV: [che lavoro deve fare? #0\_5 qual è 'sto lavoro?

SHI: || eh che:: # mh:: # sono dei- de:: #1\_2 costruttori di:: navi # tipo ||  
 INV: ah  
 SHI: || allora lui:: dovrà eh:: ehm:: raccogliere pezzi di:: legno mh |  
 {per::} per poterla fare || #  
 SHI: || gli hanno detto | di prenderne un tipo #0\_8 specifico || #  
 SHI: || va ||  
 SHI: || prende questo ||  
 SHI: || non è quello || #  
 SHI: || va ||  
 SHI: || prende questo ||  
 SHI: || è l'unico | che ha #0\_8 trovato || # .hh  
 SHI: || era l'unico:: | che:: teneva la:: #0\_5 nave #1\_0 ferma || #0\_5  
 SHI: || ah. lui. visto bello || #1\_0  
 SHI: || toglie ||  
 SHI: || non esce ||  
 SHI: || secondo te # se è là | è perché magari | c'è un:: motivo || #0\_5  
 SHI: || martello pun pun pun || #  
 SHI: || toglie {la me:} # la nave | che se ne va per i cavoli suoi ||  
 SHI: || tutti li guardano ||  
 INV: he va nel mare? [hehe  
 SHI: || [sì sì sì ||  
 SHI: || se ne va per cavoli suoi lui zitto zitto || #  
 SHI: || giacca || #  
 SHI: || cappello: ||  
 SHI: || e va #0\_8 via | perché sa | che #0\_2 ormai è licenziato dopo questa ||  
 #2\_2

Il confronto tra gli esempi (80) e (81) realizzati rispettivamente a t1 e t4 mostrano poi come la riduzione dell'incidenza di subordinate sia legata in alcuni casi a una complessificazione a livello di clausola o all'acquisizione di un lessico più specifico. Qui l'uso della parola *panettiere* porta alla riduzione della quantità di subordinate necessarie a t1.

(80) SHI: || trova: # {il: # e:il: .h #0\_5 } l'uomo | che sta facendo # il pane. | che dal furgone va dentro {a a} # al negozio? ||

(81) SHI: || e c'era giusto giusto il:: # panettiere | che stava ehm:: # scaricando il:: ehm il pane:: # dal camion || #

La traiettoria di sviluppo registrata da Shirley mostra come ai livelli più avanzati, una volta raggiunto il grado di complessità dei parlanti nativi, la variazione longitudinale può riguardare non tanto un'ulteriore complessificazione, ma piuttosto la sperimentazione di diversi stili o strategie retoriche. Nel caso di Shirley lo stile narrativo si fa più semplice e colloquiale, questo porta alla realizzazione di una minima codifica morfosintattica, con le varie parti del discorso che vengono cucite con riprese anaforiche ad hoc,

rimandi all'universo del discorso o un uso espressivo e incalzante dell'intonazione. Le AS-unit impiegate risultano più brevi e semplici o caratterizzate da frasi ellittiche e strutture pseudo-subordinanti, in ragione delle caratteristiche di questo tipo di parlato, in cui, per ottenere un maggior agio ricettivo le informazioni sono distribuite su più frasi e distribuite in modo che in prima posizione compaia l'elemento di interesse.

Nel task intervista, come per Eden, anche la produzione di Shirley a t1 equivale a quella delle studentesse italiane, sia rispetto alle misure di complessità che alla distribuzione di clausole per tipo. In prospettiva longitudinale non ci sono cambiamenti sostanziali tra t1 e t4 rispetto alle due misure di complessità, così come esemplificato in (82) e (83).

- (82) INV: [che cosa è successo il primo gi[[orno.  
 SHI: || [[.h all'inizio sono andata | || h.  
 INV: e cosa hai pensato quando dovevi andare.  
 SHI: || | h male | || #  
 SHI: || | stavo tremando. | ||  
 INV: !stavi tremando!?  
 SHI: || | sì: con tutti | ||  
 SHI: || | perché #0\_2 eravamo appena venuti | ||  
 SHI: || | una settimana dopo ave abbiamo cominciato. | ||  
 SHI: || | .h perciò vedere tutta sta gente bianca davanti a me:, || #0\_2 .h  
 SHI: || | poi all'i:nizio sono venuti tutti | perché credevano | che fossimo dall'america. | ||  
 INV: okay  
 SHI: || | visto che {hanno v crede ha} hanno sentito | che parlavo bene l'in#0\_2glese | sono venuti tutti | ||.h  
 SHI: || | appena hanno sentito dall'africa | tutti si son[o ritirati indietro | ||  
 SHI: || | ho detto | ||  
 SHI: || | guardate | che a me proprio no. .h non me ne frega niente | ||  
 SHI: || | potete starci | ||  
 SHI: || | o non starci perciò .h | ||  
 SHI: || | io infatti in seconda: ho fatto le mie cose | ||
- (83) SHI: || | abbiamo cominciato le superiori | || #  
 SHI: || | tanto per non staccarci #0\_2 | lei è andata a fare i i i i servizi sociali  
 INV: mh mh  
 SHI: e io al cattaneo | || #0\_2  
 SHI: || | e solo che sono due scuole {e e e in in} in un'unica # 1\_4 struttura | ||  
 INV: mh mh  
 SHI: || | perciò::: stavamo::: #0\_6 insieme | || #0\_8  
 SHI: || | ha fatto::: tre anni | || #0\_4  
 SHI: || | e solo che::: è stata::: bocciata #1\_8 | perché poi {l'ingle} # lei aveva perso già::: voglia | || #  
 SHI: || | aveva già venti anni | || #2\_4  
 SHI: || | sai | ||

INV: quindi era un po' grande forse per stare a scuola  
 SHI: || grande || #0\_4  
 SHI: || sì ||  
 SHI: || poi:: voleva dare una mano:: anche a mamma e papà ||  
 SHI: || perciò:: voleva già smettere per un # lavoro ||

Infine, rispetto alle telefonate di servizio, anche Shirley a t1 si discosta dalle modalità di apertura delle studentesse italiane: le sue interazioni sono brevi e complesse, le informazioni tendono ad essere impacchettate in un'unica unità. Ad esempio, in (84) dopo l'apertura del canale e i saluti, Shirley realizza una AS-unit di 2 clausole che contiene la mossa di autoidentificazione e la pre-richiesta, seguite dalla richiesta vera e propria, anche in questo caso di due informazioni molto precise, *perciò volevamo chiedere il costo dell'aereo e gli # orari*. A differenza di quanto accade per le studentesse italiane qui l'interazione non è fluida, alla densità informativa si accompagnano pause medio lunghe tra turni e tra mosse comunicative.

- (84) SHI: ((squillo))  
 AG3: moito viaggi?  
 SHI: || { .hhh } #0\_8 buonasera || #0\_5  
 AG3: buonasera  
 SHI: || allora noi siamo un # gruppo di  
 quattro # amici | che # vorremmo  
 far un viaggio # a londra || #0\_4  
 AG3: sì  
 SHI: || perciò volevamo chiedere il costo  
 [dell'aereo e gli # orari || #0\_5  
 AG3: ho [capito e: quando vorreste andare a londra

In prospettiva longitudinale Shirley non sembra ottenere cambiamenti sostanziali rispetto alle misure di complessità, rimanendo su valori superiori a quelli delle studentesse italiane. Shirley dimostra però diversi progressi sul piano interazionale, poiché già a t2 non condensa più l'entrata in merito in un'unica AS-unit come a t1, e produce diversi turni e diverse AS-unit. A differenza delle italiane però continua a realizzare una maggiore quantità di subordinate, producendo formule del tipo *allora vorrei sapere:: dei informazioni se avete xxx* (85) o *allora volevo:: sapere | se avete già {un nome di un} un titolo di [# {un} un album* (86), e dunque continuando a discostarsi sul piano pragmatico dalle studentesse italiane, che realizzano invece pre-

richieste del tipo *volevo chiederle un'informazione* come unità a sé e realizzando poi la richiesta vera e propria sotto forma di domanda.

- (85) SHI: ((squillo))  
 SH2: (libreria muratori buonasera)  
 SHI: || | buonasera | ||  
 SH2: (buonasera)  
 SHI: || | allora vorrei sapere::: dei informazioni | se avete alcuni # libri | ||  
 SH2: sì  
 SHI: || | allora Danielle Steel | ||  
 SH2: sì  
 SHI: || | il ranch | ||  
 SH2: vediamo
- (86) SHI: ((squillo))  
 SH3: ricordi mediastore buongiorno sono cristina #  
 SHI: || | buongiorno | || #  
 SHI: || | a[llora e:: | ||  
 SH3: [prego  
 SHI: || | vorrei sapere {delle::: e} e delle informazioni sul # gruppo latino americano a#ventura | || #  
 SHI: || | allora volevo:: sapere | se avete già {un nome di un} un titolo di [# {un} un album | ||  
 PAN: [°un cd°

L'esempio (87) realizzato a t4 mostra come dal punto di vista interazionale Shirley segua ormai il pattern nativo. Le informazioni sono introdotte una alla volta per favorire la graduale messa a fuoco del topic da parte dell'interlocutore. Dal punto di vista delle scelte linguistiche però sembra preferire l'uso di AS-unit più complesse. Tale scelta risulta più elaborata, ma comunque naturale e fluente.

- (87) SHI: ((squillo))  
 EP1: pronto #0\_2  
 SHI: || | pronto | || #  
 SHI: || | buonasera | || #  
 SHI: || | parlo con francesco? | || #0\_5  
 EP1: sì # sono io #0\_5  
 SHI: || | salve | || #  
 SHI: || | allora eh sono la eh #0\_5 benedetta | || #  
 SHI: || | volevo sapere delle informazioni per:: milano | ||  
 SHI: || | perché devo organizzare:: un #0\_2 viaggio | ||  
 SHI: || | allora volevo sapere | se lei eh mi poteva dare delle informazioni | su:: # cosa {vede- eh} visitare | ||  
 SHI: || | non so | || #  
 SHI: || | {gli al::: .hh} gli alberghi .hh e [quanto costano | ||  
 SHI: || | e::: qualche altro informazione che # [può essere utile | || #0\_8



#### 4.1.4. Riassunto

In § 4.1 è stata riportata l'analisi della variazione della complessità sintattica delle produzioni dei soggetti rilevata utilizzando le due misure quantitative di lunghezza e di subordinazione e le sei misure qualitative relative alla distribuzione di tipi di clausole e di subordinate. I dati mostrano variazione longitudinale e situazionale, sia a livello di gruppo che individuale. Vediamo con quali tendenze.

Longitudinalmente, le apprendenti mostrano progressi in termini di lunghezza e di distribuzione di clausole e subordinate per tipo, mentre cumulativamente in tutti i task la misura di subordinazione segna una lieve riduzione. Questo dato è in linea con quanto evidenziato in letteratura (cfr. Ipotesi di Predizione dello Sviluppo, § 2.2.2.2): a livelli più avanzati il progresso in complessità sintattica riguarda non tanto l'uso di una maggior quantità di subordinate, ma piuttosto di una maggior varietà e di una complessificazione a livello di clausola.

Situazionalmente, le studentesse italiane dimostrano variazione tra task, con quelli monologici che risultano più complessi di quelli interattivi. Anche per le apprendenti c'è variazione, ma alla prima rilevazione con una polarità opposta, le ragazze usano una sintassi tendenzialmente più complessa nei task interattivi. Nel tempo i valori di gruppo mostrano un avvicinamento ai valori delle italiane con la traiettoria di sviluppo di ciascuna misura che cambia da task a task, evidenziando come il tipo di attività svolta possa costituire un'ulteriore causa dell'andamento non lineare della complessità sintattica. Nei task interattivi il progresso è ottenuto invece per tutte le apprendenti attraverso una riduzione delle misure di complessità e un conseguente riallineamento ai valori delle compagne italiane. I dati confermano l'importanza della valutazione di questa dimensione della proficiency in diversi contesti comunicativi.

Anche individualmente c'è variazione. Per le studentesse italiane questo è evidente nei task monologici, dove una risulta più complessa. Per le apprendenti, i dati mostrano come due di loro siano più avanzate già alla prima rilevazione, nel tempo ciascuna segue un proprio andamento e ritmo, con cambiamenti più o meno lineari e più o meno repentini. Nei task monologici le apprendenti più avanzate realizzano progressi soprattutto in termini di lunghezza, le studentesse meno avanzate in termini di sviluppo della subordinazione, con un effetto di trade-off tra le due misure, accompagnato da un avvicinamento qualitativo nelle misure relative alla distribuzione di clausole e di subordinate per tipo. I dati disaggregati per soggetto confermano l'andamento non sempre lineare della complessità nel tempo, mostrando come ciò sia legato, oltre che al livello e al task considerato, anche ad un'interazione di questi fattori con la variazione individuale. Il gruppo definisce il quadro di variazione generale entro il quale poi si sviluppano i percorsi individuali di ciascun apprendente. Il confronto tra gruppo e individuo risulta dunque utile nell'elaborare profili più affidabili.

Infine, rispetto alle misure, i dati confermano come queste non siano sempre indicatori diretti della competenza, nel senso che un loro aumento non implica automaticamente progresso. Infatti i suoi valori vanno sempre considerati come il risultato di un'interazione con la variazione situazionale e individuale. Attraverso la loro combinazione e il confronto con il comportamento dei parlanti nativi e con i valori del gruppo è però possibile ottenere profili individuali di proficiency affidabili.

#### 4.2. *L'accuratezza CAF*

Vengono qui presentati i risultati relativi all'analisi dell'accuratezza effettuata calcolando la percentuale di AS-unit senza errori sul totale delle AS-unit prodotte. Si presenterà prima, per i due gruppi di soggetti, l'analisi della variazione longitudinale cumulativamente in tutti i task (§ 4.2.1), si procederà

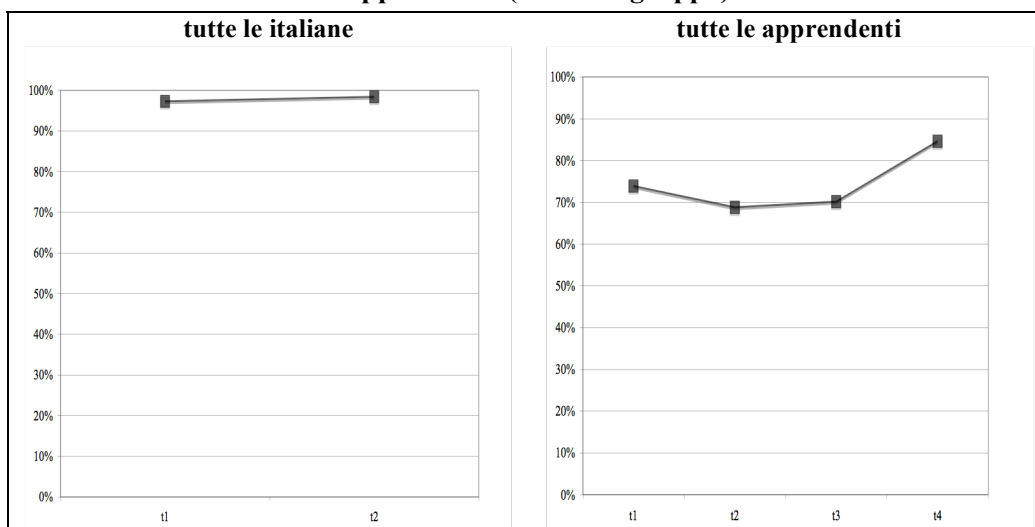
con quella della variazione situazionale tra task (§ 4.2.2), e si concluderà illustrando il percorso individuale di ogni studentessa (§ 4.2.3).

#### 4.2.1. La variazione longitudinale

Il grafico 9 mostra i valori medi dell'accuratezza per i due gruppi di soggetti relativamente a tutti i task.

Alla prima rilevazione le apprendenti, come atteso, sono meno accurate delle compagne italiane. In prospettiva longitudinale però esse ottengono un aumento dei valori di accuratezza, anche se la traiettoria di sviluppo non è lineare. Tra t1 e t3 si assiste a una lieve riduzione, seguita poi da un rapido incremento tra t3 e t4. Vedremo in § 4.2.2 a cosa è dovuta questa non linearità.

**Grafico 9. CAF – Accuratezza – Variazione longitudinale – Tutti i task – Italiane e apprendenti (valori di gruppo)**

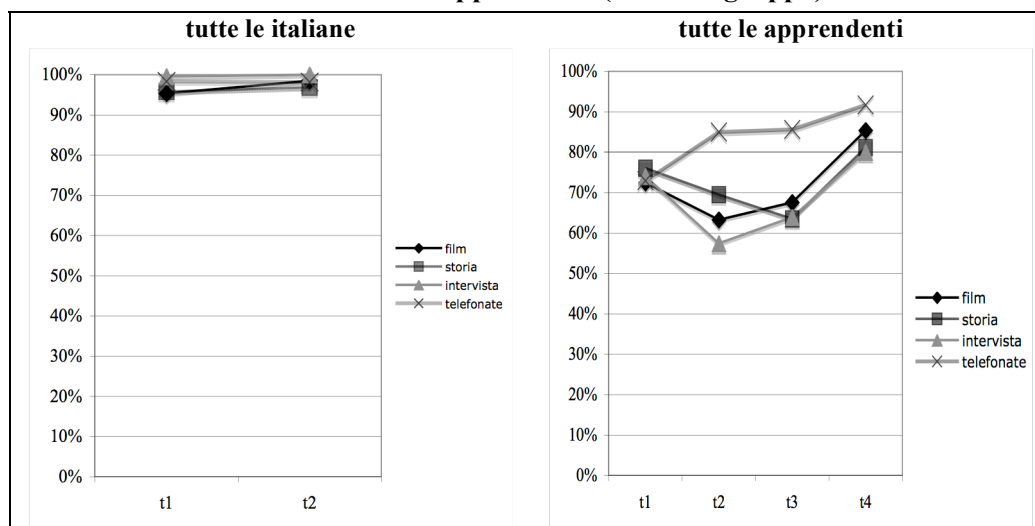


#### 4.2.2. La variazione situazionale

Il grafico 10 mostra le medie delle studentesse italiane e delle apprendenti mettendo a confronto i quattro task. I dati disaggregati per task riconfermano solo parzialmente il quadro generale descritto in § 4.2.1.

Per le studentesse italiane la variazione tra task è minima in entrambe le rilevazioni, poiché il livello di accuratezza è sempre vicino al 100%. Solo i task monologici risultano un po' meno accurati di quelli interattivi. Abbiamo già evidenziato in § 4.2.1.1 come il task racconto di un film, richiedendo ai soggetti di ricordare le sequenze video viste una sola volta e di pianificare l'esposizione, genera alcune incertezze nelle produzioni delle studentesse italiane. Questo dato è in linea con i risultati condotti nell'ambito dei modelli psicolinguistici dell'*information processing view* che mostrano come esista un legame stretto tra accuratezza e richieste del task, con effetti di trade-off (cfr. § 2.2.1.2; tabella 3 in Appendice).

**Grafico 10. CAF – Accuratezza – Variazione situazionale – Distribuzione per task – Italiane e apprendenti (valori di gruppo)**



Per le apprendenti non c'è variazione tra task a t1. In prospettiva longitudinale le telefonate si discostano notevolmente dagli altri tre task. Mentre in questi ultimi la traiettoria non è lineare, nell'apertura delle telefonate

lo è, con un graduale aumento nel tempo. Inoltre a t4 questo è il task in cui le apprendenti risultano più accurate.

I dati disaggregati per task permettono pertanto di rilevare due ragioni per la non linearità della traiettoria: l'interazione tra task e accuratezza (cfr. modelli psicolinguistici in § 2.2.1.2) e tra task e sviluppo. Rispetto alla prima ragione, per le apprendenti, come per le studentesse italiane, i due task monologici e, anche se in misura minore, il task intervista richiedono l'uso di strutture più complesse rispetto al task interattivo di apertura delle telefonate, come pronomi clitici, discorso indiretto o alcune forme di subordinazione complessa, non sempre perfettamente acquisite dalle apprendenti. Di conseguenza in questi task si realizza un effetto di trade-off tra richieste del task e accuratezza, con una conseguente traiettoria di sviluppo non lineare. Nel task relativo alle telefonate invece, dove si favorisce l'uso di poche strutture semplici, la traiettoria di sviluppo dell'accuratezza è lineare e la misura impiegata segna un progresso graduale.

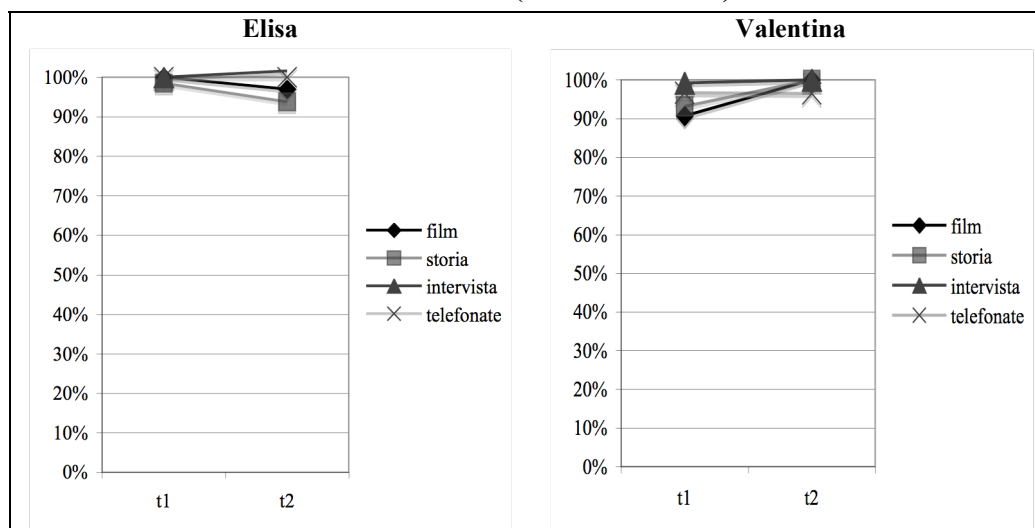
Rispetto alla seconda ragione, nei task monologici l'incremento di errori è dovuto all'aumento delle sperimentazioni di forme complesse o alla ristrutturazione di forme già conosciute (così come verrà esemplificato in dettaglio in § 4.2.3). Questo dato è in linea con i risultati di altre ricerche acquisizionali. Già in uno studio del 1983, Larsen-Freeman dimostra come l'accuratezza tenda a diminuire passando dal livello intermedio a quello più avanzato. L'apprendente in una fase iniziale non produce la struttura, e molto spesso ne evita anche i contesti d'uso, mostrandosi dunque poco evoluto ma accurato. In una successiva inizia a sperimentarne l'uso, creandosi più contesti, impiegandola in modo corretto in alcuni casi e errato in altri o generalizzandola a contesti in cui non sarebbe richiesta, risultando pertanto più avanzato da un punto di vista interlinguistico ma meno accurato. Infine la struttura viene impiegata in modo più vicino a quello nativo. In altre parole l'accuratezza è determinata dall'interazione tra elementi quali: grado di accuratezza della rappresentazione stessa, forza di eventuali rappresentazioni in competizione tra loro e grado di automatizzazione, così come suggerito in diversi studi

acquisizionali (cfr. § 2.1.2; § 2.2.2.2). Questo dato permette di fare un'ulteriore riflessione: accuratezza e sviluppo interlinguistico rappresentano due costrutti diversi, pertanto un apprendente può essere accurato ma poco evoluto, o più evoluto ma poco accurato. Poiché l'accuratezza interagisce con le ristrutturazioni che ricorrono nelle fasi in cui si stanno acquisendo nuove strutture, lo sviluppo in questa dimensione può seguire andamenti a U, con momentanei regressi, seguiti da progressi (cfr. Wolfe-Quintero, 1998; § 2.2.2.2).

#### 4.2.3. La variazione individuale

Le studentesse italiane (grafico 11) pur riconfermano il quadro delineato dai dati di gruppo dimostrano una leggera variabilità a livello individuale con Valentina che risulta anche un po' meno accurata della compagna.

**Grafico 11. CAF – Accuratezza – Variazione individuale – Distribuzione per task – Italiane (valori individuo)**



Abbiamo visto in § 4.2.2 come questo sia legato alle richieste dei due task: dovendo ricostruire la trama di una storia le studentesse non sempre riescono a prestare attenzione alla pianificazione del discorso e al controllo della forma

contemporaneamente, in alcuni casi pertanto producono alcune imperfezioni (88-90). In altri casi invece le incertezze sono legate all'uso di forme dell'italiano popolare (91). L'analisi qualitativa degli errori mostra come questi siano oltre che ridotti in numero, 5 casi in tutto, limitati ad alcune strutture specifiche. In tre casi si tratta dell'uso o dell'accordo del pronome clitico, Valentina ad esempio utilizza frequentemente il pronome clitico *lo* per riferirsi alla rana protagonista nella storia, accordando quindi il pronome per senso, la rana protagonista è probabilmente di sesso maschile, e non grammaticalmente, la rana è un sostantivo femminile. Elisa in un caso non realizza invece la ripresa, così come esemplificato in (88).

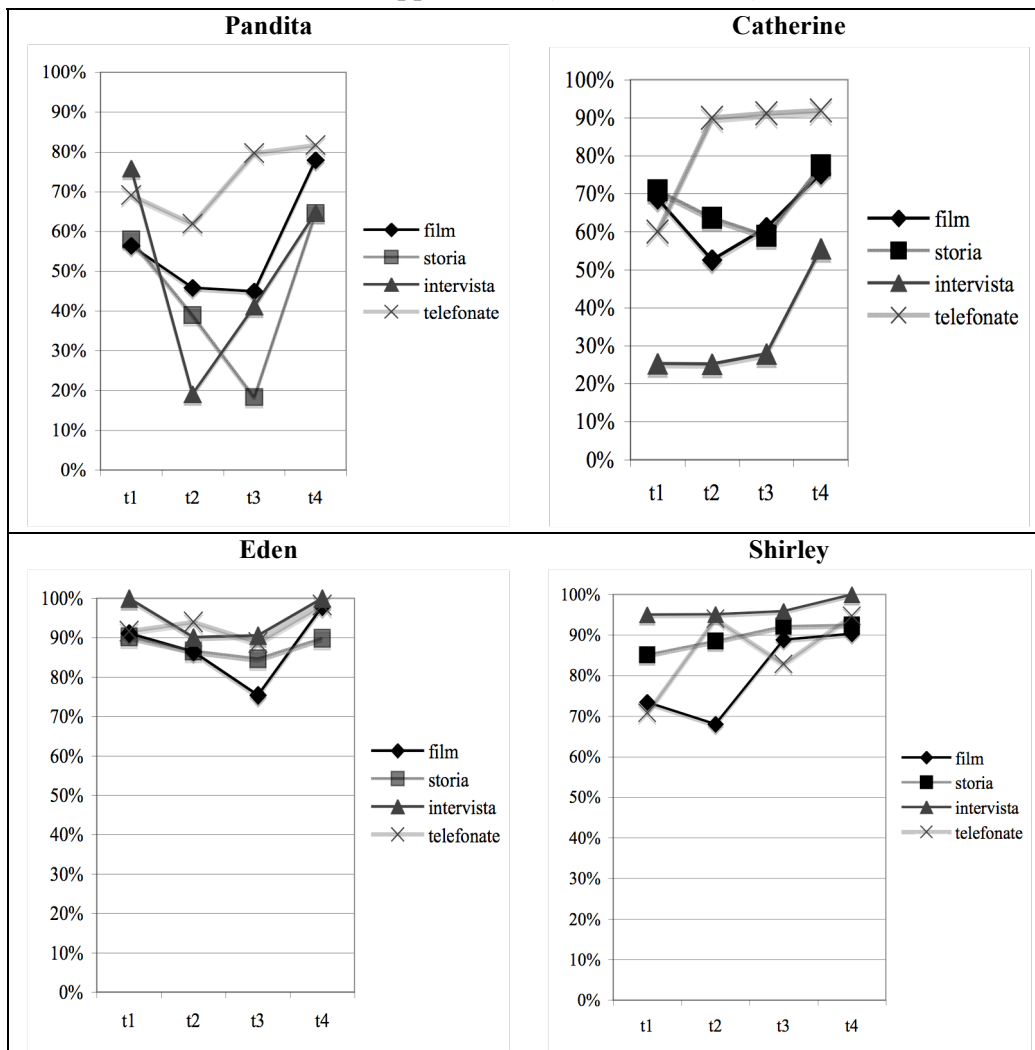
- (88) ELI: || | solo che:: ehm # durante una notte la pantera rosa | che era fuori al  
freddo | aveva bisogno | di riscaldarsi | ||# .hhh  
ELI: || | e::: {va:::} ruba {un pezzo di legno} #0\_3 insomma del fuoco:: dalla  
casa del cacciatore # || 0\_3.hhh  
ELI: || | il cane corre dietro | per {pre} riprendendosi il bastoncino {con} con il  
fuoco? | || #0\_8

Le altre occorrenze riguardano un caso di accordo del participio passato, così come esemplificato in (89) realizzato da Elisa, due casi di uso errato della forma di subordinazione, così come in (90) e (91), realizzati da Valentina.

- (89) VAL: || | hh solo che::: appunto vede questo::: #0\_2 cioè Charlie Chaplin | {che:  
dice # cioè} che: h dice | ||  
VAL: || | {pre} non prendete lei | ||  
VAL: || | ma prendete me | perché son stata io | || #0\_4
- (90) VAL: || | quindi la polizia: | vedendo | che non paga #0\_4 | {lo a} lo prende | da  
portare in galera | ||
- (91) VAL: || | sembra | che si sta facendo il bagno qua £h[h£ .h sullo stagno {col cane  
in te- e} col cane in testa | ||

Per le apprendenti i percorsi individuali dimostrano notevole variabilità. Il grafico 12 raffina il quadro appena tracciato con i dati di gruppo per la variabilità longitudinale e situazionale.

**Grafico 12. CAF – Accuratezza – Variazione individuale – Distribuzione per task – Apprendenti (valori individuo)**



Pandita e Catherine sono meno accurate delle altre due compagne, con una certa variazione tra task, così come per le compagne italiane, anche nel loro caso il task interattivo di apertura delle telefonate risulta più accurato degli altri tre. In prospettiva longitudinale le due studentesse confermano il quadro di variazione identificato per il gruppo: sono entrambe più accurate in tutti i task,



mentre nei due task monologici, e per Pandita anche nel task interattivo intervista, le traiettorie di sviluppo non sono lineari, mentre lo sono nel task interattivo di apertura telefonate, e nel task intervista per Catherine.

Eden e Shirley, già piuttosto avanzate fin da t1, diversamente dalle altre compagne italiane mostrano meno variazione tra task e con tendenze opposte, il task di apertura delle telefonate risulta meno accurato degli altri. In prospettiva longitudinale, Eden segue lo stesso percorso di sviluppo di Pandita e Catherine, per Shirley invece l'accuratezza cresce in tutti i task. Il livello iniziale può dunque costituire una terza causa di variazione che va ad aggiungersi a quelle identificate in § 4.2.1 e § 4.2.2, cioè lo sviluppo interlinguistico e il task. Vediamolo ora in dettaglio attraverso esempi dal corpus.

Per Pandita a t1, così come per le italiane, i task interattivi sono un po' più accurati di quelli monologici e in prospettiva longitudinale ci sono differenze tra task nelle traiettorie di sviluppo. Vediamone le tendenze.

Nei task monologici gli errori commessi da Pandita sono riconducibili ad alcune strutture in particolare, così come esemplificato dagli estratti che seguono, tratti dal task film. Oltre ad alcune incertezze lessicali, legate soprattutto alla scelta delle preposizioni, si ritrovano errori nell'uso dei clitici (92), nella maggior parte delle occorrenze assenti, con qualche sperimentazione di *lo*, dell'articolo, anch'esso assente in buona parte dei contesti obbligatori (93), del passato prossimo, sia per quel che riguarda la scelta dell'ausiliare che l'accordo del participio passato, delle forme riflessive, con l'omissione della particella *si* nella maggior parte dei casi, e infine la scelta delle preposizioni (94).

- (92) PAN: || | rubato £io il pane | ||  
PAN: || | ce l'ho io pa[nef | ||

- (93) PAN: || | allora lui dà: # suo posta alla ragazza? | ||

- (94) PAN: || | e::: la ragazza #0\_3 cade sopra: h #0\_7 il uomo .h | che l- era andato a  
lavorare {in: °n na- # } in nave per la°  
INV: .h ne nel cantiere  
PAN: cantiere | ||

In prospettiva longitudinale l'andamento della traiettoria di sviluppo non è lineare, dopo una riduzione del grado di accuratezza a t2 e a t3, a t4 Pandita raggiunge valori ben superiori a quelli iniziali. Se a t2 e t3 Pandita risulta quantitativamente meno accurata, l'analisi qualitativa del tipo di errori mostra come tale riduzione sia legata alla sperimentazione di nuove strutture o alla ristrutturazione di alcune forme. È il caso, ad esempio, dei verbi riflessivi. A t2 l'uso è corretto in alcuni contesti (95), in altri si assiste a generalizzazioni, come esemplificato in (96) e in (97).

(95) PAN: || | £hh£ # il cane si accorge | che:: #0\_8 {e un:} non è una statuo. | è un  
#1\_0 un liono.# riale.||

(96) PAN: || | e si dorme | ||

(97) PAN: || | {e::} e pink panther.# {e si::} si sente | che arriva padrone. | || #

In modo analogo procede l'uso di AS-unit con dipendenti complesse. Nell'esempio (98) realizzato a t3 notiamo come la sperimentazione di relative complesse sia accompagnato da incertezze grammaticali e produca forme scorrette.

(98) PAN: || | però un certo punto {il il liono} # il pink panther f\*ha sentito la  
voce::\*f de:l {s:} qualcuno | ha {sp} sparato | ||

Infine, Il confronto tra gli esempi (99) e (100), realizzati a t2 e t3 rispettivamente, illustra chiaramente come l'acquisizione di forme più complesse generi un andamento non lineare della misura di accuratezza.

(99) PAN: || | { e poi f\*questa\*f statua era:::}? | invece lui è:#0\_4 come una statuo. |  
||

(100) PAN: || | mette sopra la sua testa | # facendo finta che | se come se lui è un  
albero. | || #0\_6

A t4 la produzione di Pandita risulta in generale più accurata rispetto alle strutture sopra descritte, come esemplificato in (101)

(101) PAN: || | prima li ha svegliati {il} #0\_2 il ragazzo | ||  
PAN: || | e poi ha fatto svegliare la ragazza | ||  
PAN: || | {gli ha e} gli ha detto | di scappare:: | ||  
INV: mh mh

PAN: || | pertanto:: # in quel momento c'era anche il poliziotto | che diceva | di svegliarla | ||

Nel task intervista, Pandita a t1 risulta essere più accurata che in quelli monologici, con valori vicini all'80%. Le sue interazioni sono realizzate per frasi brevi e semplici, in questa prima fase Pandita è poco evoluta ma sufficientemente accurata. In prospettiva longitudinale, così come per i due task visti sopra, si assiste a una rapida riduzione della misura di accuratezza, con un livello addirittura inferiore al 20% tra t2 e t3, seguita da un ulteriore cambiamento in positivo a t4. Nonostante tale riduzione, nel tempo la conversazione risulta man mano più fluida, fatta di AS-unit più lunghe e complesse. Proprio questo sviluppo qualitativo dell'interazione offre da un lato, più contesti per strutture più complesse, dall'altro maggiori occasioni di commettere errori, come evidente dal confronto tra le produzioni a t1 (102) e le produzioni a t2 (103) e a t3 (104-106). In questa seconda fase, Pandita è quindi più evoluta ma meno accurata. A t4 si confermano i progressi qualitativi, a cui segue un aumento quantitativo del livello di accuratezza rispetto a t1.

- (102) PAN: || | f\*ho detto\*f ? | ||  
 PAN: || | bonporto? | ||  
 [...]  
 PAN: || | allora son salita | ||  
 PAN: || | h f\*son venuta\*f qua. | #  
 PAN: | perché ho fatto questa scuola. | ||  
 INV: sì. #  
 PAN: || | e son entrata. | || #  
 PAN: || | io avevo solo foglio scritto.#0\_6 terza effe. | || #  
 PAN: || | f\*ho fatto vedere.\*f | # dov'è £terza?[effe.£] | ||  
 INV: [£hh£]  
 PAN: || | [mi hanno detto | ||  
 PAN: || | su. su. su. | ||  
 PAN: || | sono andata avanti. | ||  
 PAN: || | son salita. | || #0\_2  
 PAN: || | c'era prof paola.# di italiano. | che aspettava. | || #
- (103) PAN: || | però,#0\_2 secondo me è stato bene.|# perdere un anno. | || #  
 PAN: || | quello terza media,# mi ha aiutato molto. | per imparare lingua. | || #  
 [...]  
 PAN: || | {se} magari se. # andavo. # subito a superiore. | era molto difficile per me. | || #  
 INV: perché.#  
 PAN: || | perché {tes com} la chimica | che faccio adesso | è molto difficile. | || #
- (104) PAN: || | quindi {non} lei non ha mai eh bisogno di dirmi | che stu:dio | ||

(105) PAN: || | mio padre voleva | che mh magari lui faccia una scuola | così  
imparasse bene lingua .hh | ||

(106) PAN: || | l'unico posto # in cui # con la quale si può lavorare senza anche scuola  
||

Anche il task interattivo di apertura delle telefonate è a t1 più accurato di quelli monologici, ma in prospettiva longitudinale ci sono differenze rispetto alla traiettoria di sviluppo che, contrariamente a quanto avviene negli altri task, qui è lineare.

A t1 le AS-unit di richiesta realizzate da Pandita sono nella maggior parte dei casi poco standard nella forma (107).

(107) PAN: || | {h} vorrei sapere informazioni sul telefono | || #0\_5

In prospettiva longitudinale i cambiamenti seguono una traiettoria lineare. A t2 forme ora più accurate (108) si alternano a sperimentazioni che portano a errori formali (109), a t3 (110-111) e a t4 (112) Pandita evidenzia progressi in due direzioni, da un lato usa forme accurate, dall'altro differenzia le modalità di richiesta.

(108) PAN: || | io volevo chiedere # un'informazione su un libro | ||

(109) PAN: || | perché io volevo comprare un libro per mio amico. | #0\_5 che:: lei  
piace libri di romanzo. | || #2\_0

(110) PAN: || | posso chiedere {un'infor} un'informazione | ||

(111) PAN: || | io vorrei sapere | se magari lei sa # un titolo di un cd | || #

(112) PAN: || | mi chiamo pandita, | ||

EA1: sì::

PAN: || | e vorrei sapere di alcuni divudi | || #0\_8

A t4 Pandita ottiene valori di accuratezza molto alti, nonostante permangano ancora alcune inappropriatezze formali di tipo morfosintattico in due delle formule di richiesta impiegate, come esemplificato in (113) e (114).

(113) PAN: || | vorrei sapere su informazioni su::: # un viaggio a milano | || #1\_2

(114) PAN: || | vorrei sapere delli costi {sul} # di alcuni film | || #1\_5

Catherine mostra un quadro di variazione vicino a quello di Pandita. A t1 è meno accurata nell'intervista, ma più accurata negli altri task. In prospettiva

longitudinale anche per lei c'è variazione nelle traiettorie di sviluppo tra task monologici e interattivi, con differenze tra i due anche più significative rispetto alla compagna. Vediamo alcuni esempi dal corpus.

Nei task monologici Catherine non si discosta molto dal percorso di Pandita, nonostante a t1 risulti un po' più accurata. Anche nel suo caso gli errori sono limitati ad alcune strutture: l'uso delle preposizioni, l'accordo del participio passato e l'uso dell'ausiliare nel passato prossimo, così come evidente dall'estratto (115) realizzato nel task film a t1.

(115) CAT: || con il signore | invece {di sedere su} # di sedere sulla sedia | (si) è  
 {sedu} è andata a sedere {su una donna} # sopra una donna || .hh  
 INV ah ah?  
 CAT: || quella donna {ha:: uff} ha spinto || #  
 [..]  
 CAT: || a certo punto seduto su una sedia ||  
 CAT: || è venuto ancora || #  
 CAT: || spinto || #  
 CAT: || quella donna {ha:: uff} ha spinto || #

In prospettiva longitudinale, anche lei, come Pandita, segue una traiettoria di sviluppo non lineare. Anche in questo caso tra t1 e t3 la riduzione è legata alla sperimentazione di forme più complesse o alla ristrutturazione di forme note, come esemplificato in (116) e (117) realizzati a t3 nel task film, dove gli errori sono legati al tentativo di utilizzare subordinate più complesse.

(116) CAT: || però::: il rano grande #1\_2 ha salto #  
 CAT: ha salito su #0\_2  
 CAT: seguendo mh | per cacciare il piccolo rano || #0\_2  
 (117) CAT: || mi sembra | che {sta-} cioè col nervoso sta dicendo | di non seguirli:::  
 # l'indicazione | dove vanno | #0\_8 di prendere::: la sua parte::: || #0\_6

A t4 lo sviluppo qualitativo è accompagnato da un livello di accuratezza globale maggiore rispetto a t1, così come esemplificato in (118).

(118) CAT: || {è andato a cercare un} # è andato a cercare un lavoro da un uomo || #  
 .hhh  
 CAT: || pero lui non lo ha preso ||  
 CAT: || {l'ha} # l'ha scritto una lettera #0\_5 | per andare a presentare l'altra  
 azienda | {se::: # la::: # se::: #} se la prendono || .hhh  
 CAT: || lui andato a fare a presente un'azienda || #  
 CAT: || l'han preso || #0\_5

Nei task interattivi Catherine segue traiettorie di sviluppo più lineare rispetto alla compagna. Nell'intervista tra t1 e t3 non ci sono variazioni sostanziali nella misura di accuratezza, ma a t4 si ottengono progressi notevoli. Come esemplificato dagli estratti in (119) e (120), se la misura di accuratezza non registra cambiamenti sostanziali tra t1 e t3, l'analisi qualitativa mostra evoluzioni qualitative dell'interlingua, con la sperimentazione ad esempio di forme più complesse, che porta a imprecisioni formali.

- (119) CAT: || diciottanni che #0\_3 sono qui || #0\_4  
 INV: !caspita! #0\_3 e tu #0\_3 quando eri in ghana con chi stavi? #0\_3  
 CAT: || con mia zia || #0\_4  
 INV: con tua zia #0\_9 .hhh #0\_5  
 INV: e come è stato rit- #0\_4 venire in italia e stare di nuovo con #0\_3 con i tuoi genitori #2\_1  
 CAT: || niente || #0\_8
- (120) CAT: || allora ero molto contenta | di vivere coi miei | perché: là abitavo con la mia zia | || #  
 CAT: || e:: # ero molto contenta | quando venivo qua | || #  
 CAT: || almeno se non c'è nessuno | che conoscevo | {av-} i miei # e:h {co-conos-} conoscevano | ||

A t4 invece i progressi interlinguistici sono accompagnati da un incremento dei valori della misura di accuratezza, così come esemplificato in (121).

- (121) CAT: || 18, 19, il mio voto è quello ||  
 CAT: || {ma io non ho} all'inizio non ho mai accettato un voto così ||  
 CAT: || perché se io inissio a accettare un voto così | poi dopo alla fine, | quando mi laureo | avrò un voto bassissimo ||  
 INV: certo!  
 CAT: || allora io # cioè # rifiuto ||  
 CAT: || poi riprovo a farlo ||  
 CAT: || e poi piano piano, | essendo che l'università è primo anno | io ci sono alcuni termini che non conosco, allora ||  
 CAT: || se conosco poco | provo a far l'esame ||

In modo analogo, per il task di apertura delle telefonate, Catherine registra progressi lineari. Se a t1, l'apprendente utilizza una sola forma di richiesta caratterizzata da incorrettezze formali (122-123), nelle rilevazioni successive non è solo più accurata, ma usa anche una maggior varietà di formule, così come esemplificato da (124) e (125-126) realizzati rispettivamente a t2 e t4.

- (122) CAT: || e:: vorrei avere informazione sul nek trecento trentotto || #0\_8
- (123) CAT: || {e:} vorrei avere informazioni ||
- (124) CAT: || vorrei chieder una cosa #  
ST6: °sui libr°  
CAT: e:: sui libri ||
- (125) CAT: || allora #0\_5 {volevo} # volevo chiederti un'informazione || #1\_0
- (126) CAT: || [vorrei avere alcune informazioni su un viaggio ||

Vediamo ora i percorsi individuali di Eden e Shirley, le due apprendenti più avanzate alla prima rilevazione. Nel loro caso i progressi sono più lineari, sebbene vi siano alcune differenze a livello individuale. Vediamo con quali tendenze.

Alla prima rilevazione Eden ottiene nei task monologici e nell'intervista livelli di accuratezza molto alti, dimostrando così di essere già piuttosto avanzata. In prospettiva longitudinale le sue traiettorie di sviluppo sono simili a quelle di Pandita e Catherine, sebbene la portata del cambiamento sia quantitativamente meno significativa. A t1 gli errori di Eden riguardano le preposizioni, l'uso dei clitici e alcune forme di subordinazione, come nel caso dell'uso del discorso indiretto (127) o di relative indirette (128).

- (127) EDE: || { f\*cioè } # insomma {la d-} \*f charlie chaplin la difende .h | dicendo |  
che sono stato io | che {ho ruba:} ho rubato questo [pane qua] ||
- (128) EDE: || e gli scrive una lettera | dove: può andare a lavorare ||

Lo sviluppo longitudinale viene realizzato anche nel suo caso attraverso un'iniziale riduzione dell'accuratezza: a t2 e t3 la ristrutturazione di forme complesse provoca maggiori imperfezioni. È il caso dell'accordo del passato prossimo con ripresa del clitico: se a t1 Eden tende ad esplicitare l'oggetto, a t3 utilizza il pronome clitico, generando però incorrettezze rispetto all'accordo del participio passato, così come nell'esempio riportato in (129), dove il participio dovrebbe essere al femminile, perché riferito a una bomba.

- (129) EDE: || allora il cane invece l'ha portato # | mentre era svenuto il padrone || #  
EDE: || cioè il cane l'ha appoggiato vicino a lui || #

A t4 la produzione di Eden risulta più accurata, ormai ai livelli delle studentesse italiane, così come esemplificato in (130)

- (130) EDE: || quando arriva lì: | gli dà quella lettera lì ||  
 EDE: || e lo accettano ||  
 INV: mh mh #0\_2  
 EDE: || poi {si::} subito comincia a lavorare || #  
 EDE: || gli dice | di trovare una tavoletta || #  
 EDE: || {cioè gli fa vedere come::} # gli dice | di cercare un'altra come questa |  
 || #0\_8  
 EDE: || lui va a prendere:: quella hehehehef | {dove} dov'è #0\_8 appoggiata  
 la barca no ||

Anche Shirley è già piuttosto accurata fin da t1 e come per le altre apprendenti gli errori sono legati ad alcune strutture specifiche: l'accordo dei clitici (131), l'accordo del participio passato nel passato prossimo, in particolare in contesti in cui ricorre il clitico (132), la formazione di alcune AS-unit complesse al congiuntivo (133).

- (131) SHI: || appena::? # l'ha preso. | {c'è} c'era la signora? .h | che le stava  
 fissando::, ||  
 (132) SHI: || ne ha trovati: #1\_0 una? ||  
 (133) SHI: || ma serviva? | .h a tener ferma la:: b- barca | per non partir {da:: .h}  
 #0\_6 da solo. ||

Si differenzia però dalle compagne rispetto alla traiettoria di sviluppo longitudinale, nel suo caso lineare anche nei task monologici e nell'intervista. L'analisi permette di osservare come tale aumento sia il risultato, da un lato di una narrazione più breve e semplice, dall'altro dall'uso più accurato di varie strutture linguistiche, come illustrato dall'evoluzione che intercorre tra l'esempio (134) prodotto a t2 e (135) a t3 nel task film.

- (134) SHI: || che:: lui ha fatto di tutto | {per far m} di sostì .hhh tuire questo cane  
 qua::, ||  
 SHI: || ha fatto di tutto | che così il pa # drone lo mandasse via. | così lui  
 restava in .hhh casa. ||  
 [...]  
 SHI: || ha fatto finta | che {l'av} l'aveva mangiato il cane. ||  
 (135) SHI: || e::: la pantera? #0\_4 {e:::} ha sempre fatto di tutto. | di modo che::: {il  
 ca} il cacciatore s::i::: incavoli con e::: il cane. || #0\_2



Nel task di apertura delle telefonate, sia Eden che Shirley seguono traiettorie di sviluppo abbastanza lineari.

Per Eden non ci sono cambiamenti molto significativi tra t1 e t4, poiché il livello di accuratezza è piuttosto alto fin dalla prima rilevazione. Se a t1 e t2 Eden produce diverse occorrenze di formule di richiesta formalmente scorrette, come quella riportata in (136), a t3 (137) e a t4 (138) l'apprendente realizza richieste attraverso formule più varie e accurate. A t4 raggiunge un livello di accuratezza molto vicino al 100%. L'unica AS-unit contenente un errore è esemplificata in (139). In questo caso si tratta di una AS-unit di richiesta piuttosto complessa. L'imprecisione qui riguarda la scelta del connettore che introduce la subordinata implicita.

(136) EDE: || volevo chiederti informazione sui telefonini. || #0\_2

(137) EDE: || allora volevo chiederle alcune informazioni:: {per #0\_8 per una::} per un # corso | che stiamo facendo ||

(138) EDE: || noi:: s:: volevamo chiedere un'informazione per una gita | che dobbiamo organizzare || #1\_0

(139) EDE: || allora io avrei bisogno di una mano:: mh # | da regalare un film giallo a una signora || #0\_6

Anche Shirley a t1 utilizza una sola formula di richiesta scorretta dal punto formale, come esemplificato in (140). La ripetizione di questo stesso errore incide in modo significativo sul livello di accuratezza globale. Tra t1 e t2 l'apprendente ottiene una repentina crescita, superando il 90% (141).

(140) SHI: || allora vorrei sapere {delle} dei: informazioni sul sul telefono nek trecentotrentotto || #

(141) SHI: || a:: allora vorrei sapere delle informazioni su un libro | da consigliare ad una mia amica. || #

A t3 si assiste a una lieve riduzione: così come per Eden gli errori sono dovuti a tentativi di produrre AS-unit di richiesta più elaborate, così come in (142) e (143).

(142) SHI: || volevo delle # informazioni e: | perché dovrei re # galare un cidì per una mia amica # di #0\_4 quarant'anni || #

(143) SHI: || | allora e:: volevo sapere delle informazioni a ri # guardo del #0\_4  
gruppo a # ventura | || #

A t4 si riconferma il generale progresso in termini di accuratezza, così come esemplificato in (144).

(144) SHI: || | allora # eh:: volevo sapere delle informazioni # | perché # eh:: dovrei  
organizzare un # vi::aggio a:: # milano | || # .hh  
SHI: || | allora volevo sapere | se lei mi poteva dare {dei::: in in} delle eh  
informazioni utili | || #0\_8

Ricapitolando, i miei dati dunque mostrano come il livello di accuratezza e l'andamento della sua traiettoria di sviluppo siano legati a tre elementi: il tipo di task, le interazioni con lo sviluppo interlinguistico, e il livello iniziale. I task interattivi tendono ad essere più accurati di quelli monologici e soprattutto il task di apertura delle telefonate ha una traiettoria di sviluppo lineare per tutte le apprendenti. I task monologici seguono traiettorie di sviluppo con andamenti che si differenziano sulla base del livello iniziale: Shirley, più accurata a t1, segue traiettorie di sviluppo lineare, le altre apprendenti seguono una traiettoria di sviluppo dell'accuratezza con un andamento ad U, effetto dell'interazione tra accuratezza e sviluppo interlinguistico.

La misura di accuratezza scelta si conferma dunque un buon indicatore globale dello sviluppo di questa dimensione, sensibile alla variazione longitudinale e situazionale segnata dal gruppo e dall'individuo. I dati evidenziano anche però come a livello avanzato non sia opportuno fare generalizzazioni sulle competenze in base a un unico risultato, poiché un aumento o una diminuzione non corrisponde automaticamente a un progresso o un regresso.

#### 4.2.4. Riassunto

In questa sezione ho riportato l'analisi della variazione dell'accuratezza misurata attraverso la misura quantitativa percentuale di AS-unit senza errori

sul totale delle AS-unit. I dati mostrano variazione longitudinale e situazionale, sia a livello di gruppo che individuale. Vediamo con quali tendenze.

Longitudinalmente, le quattro apprendenti tra t1 e t4 dimostrano progresso, le traiettorie di sviluppo non sono lineari e hanno un andamento a U, determinato dall'interazione tra accuratezza e sviluppo interlinguistico. Questo dato è in sintonia con quanto osservato in altre ricerche acquisizionali che mostrano come gli apprendenti più avanzati, poiché sperimentano forme più complesse, tendano a produrre più errori (cfr. Larsen-Freeman, 1983; Skehan, 1998; Wolfe-Quintero, 1998).

Situazionalmente, per le italiane c'è variazione tra task, i task interattivi sono più accurati che quelli monologici. Questo dato conferma quanto evidenziato dalla letteratura sui modelli psicolinguistici. Diversi task pongono diverse richieste al parlante: i task monologici richiedono sia di memorizzare una serie di eventi, che pianificarne la narrazione, generando quindi un effetto di trade-off con l'accuratezza, i task interattivi, richiedendo l'uso di formule o strutture poco complesse permettono invece una maggior attenzione alla forma. Le apprendenti dimostrano meno variazione delle compagne a t1, ma nel tempo le traiettorie di sviluppo dell'accuratezza si differenziano sulla base del task, con il task di apertura delle telefonate che segna incrementi lineari e quelli monologici e l'intervista che seguono un andamento a U

Anche individualmente c'è variazione. Due delle apprendenti sono meno accurate a t1 delle compagne e l'accuratezza nel tempo segue traiettorie più o meno lineari in relazione al livello iniziale. Se nel task apertura delle telefonate tutte le apprendenti ottengono aumenti lineari dell'accuratezza, nei task monologici e nell'intervista si registrano alcune differenze. Per tre delle apprendenti questi task disegnano traiettorie non lineari, mentre per una delle apprendenti, tra le più avanzate del gruppo a t1, i progressi sono sempre lineari.

Infine, rispetto alla misura selezionata, i miei dati mostrano come l'andamento delle traiettorie di sviluppo sia legata a tre elementi: task considerato, sviluppo interlinguistico e livello iniziale. Pertanto, così come osservato per la complessità, la percentuale di AS-unit corrette sul totale delle

AS-unit non deve essere intesa come misura diretta della proficiency linguistica, poiché un aumento o una diminuzione non sempre corrisponde automaticamente a un miglioramento. Se da un lato quindi è importante valutare l'accuratezza in diversi contesti, dall'altro per avere profili più affidabili sarebbe utile, oltre che osservare l'accuratezza in diversi task, combinare la misura selezionata con indicatori capaci di rilevare lo sviluppo interlinguistico. In questo senso potrebbe essere utile rilevare la presenza di alcune strutture con valore diagnostico rispetto agli stadi dell'interlingua.

#### 4.3. *La fluenza CAF*

Di seguito si presentano i risultati relativi all'analisi della fluenza effettuata attraverso l'applicazione di due misure analitiche: il numero di pause superiori a #0\_5, e la percentuale dei fenomeni di esitazione, entrambi calcolati sul totale delle AS-unit. Si ricorda che nel caso di queste misure il progresso è realizzato attraverso una riduzione dei valori, poiché viene considerato più fluente il parlante che produce un numero minore di pause lunghe e di fenomeni di esitazione. Si inizierà presentando per i due gruppi l'analisi della variazione longitudinale cumulativamente in tutti i task (§ 4.3.1), si procederà con quella della variazione situazionale tra task (§ 4.3.2), e si concluderà illustrando il percorso individuale di ogni studentessa (§ 4.3.3).

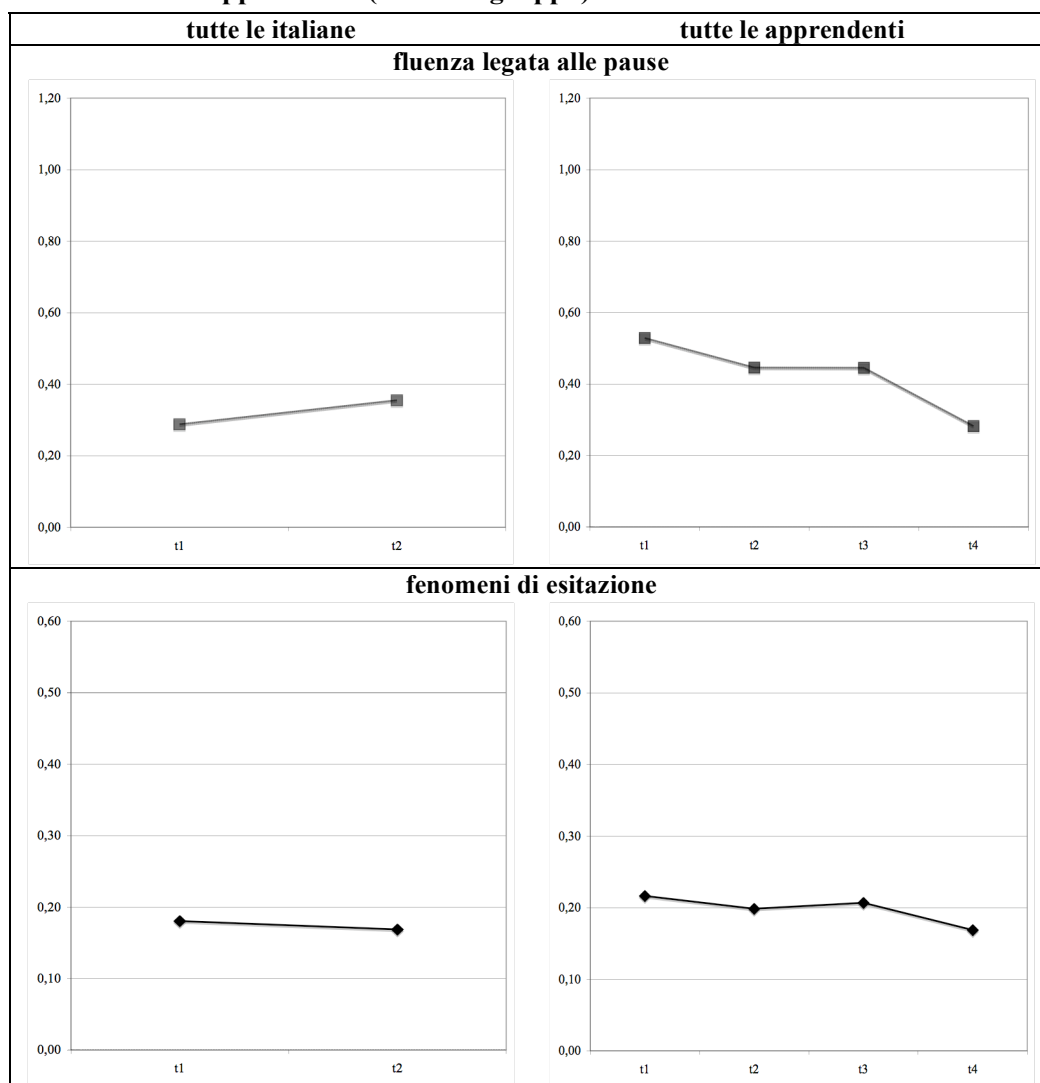
Nell'analisi di questa variabile, nel task intervista sono stati considerati solo i primi cinque minuti di conversazione in ogni rilevazione. Le conversazioni di ciascuna ragazza infatti variano notevolmente per durata, e questo task, al contrario degli altri, non ha delimitazioni ben precise. Per arrivare a un confronto più preciso è stato quindi necessario uniformare il campione, e cinque minuti sono stati considerati un tempo sufficiente per ottenere risultati validi (cfr. per es. Foster & Skehan, 1996).

#### 4.3.1. La variazione longitudinale

Il grafico 13 mostra la media dei valori ottenuti tra t1 e t4 dall'applicazione delle due misure di fluenza alle produzioni dei due gruppi delle studentesse italiane e delle apprendenti, globalmente in tutti i task.

A t1 le apprendenti sono complessivamente meno fluenti delle italiane rispetto al numero delle pause, ma non per numero di esitazioni. Tra t1 e t4 si allineano ai valori delle compagne italiane anche nella prima misura.

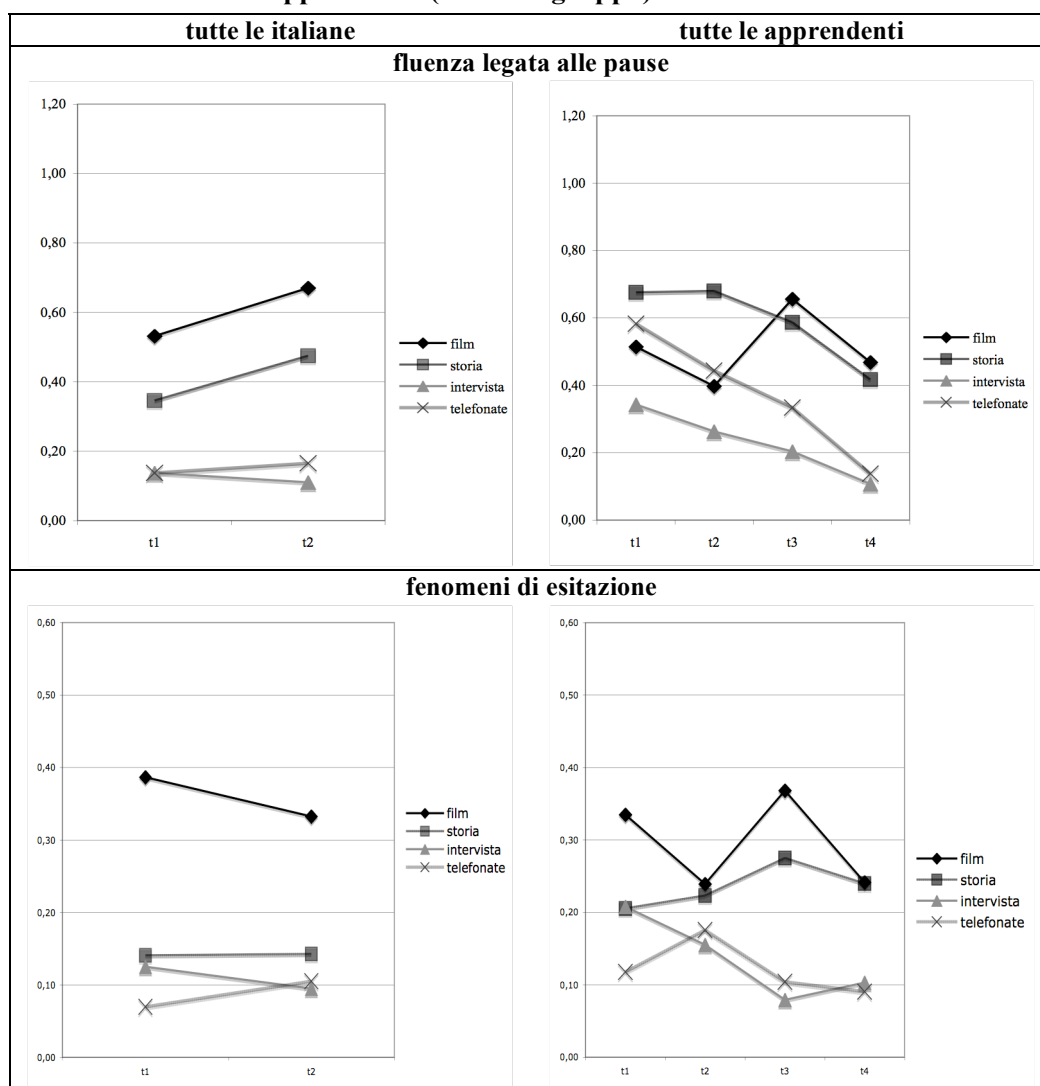
**Grafico 13. CAF – Fluenza – Variazione longitudinale – Tutti i task – Italiane e apprendenti (valori di gruppo) – Pause ed esitazioni**



#### 4.3.2. La variazione situazionale

Il grafico 14 mostra la media dei valori ottenuti dai due gruppi di soggetti nelle due misure di fluenza disaggregati per task. Questi valori rivelano notevole variazione situazionale, sia tra le italiane sia tra le apprendenti e riguardo sia alle pause sia alle esitazioni.

**Grafico 14. CAF – Fluenza – Variazione situazionale – Distribuzione per task – Italiane e apprendenti (valori di gruppo) – Pause ed esitazioni**



Per le studentesse italiane i task monologici risultano meno fluenti degli interattivi. Questo dato è in linea con i modelli psicolinguistici (§ 2.2.2.2), secondo i quali le richieste comunicative dei diversi task favoriscono diverse dimensioni della produzione, con i task interattivi che richiedono la produzione di turni rapidi e brevi, con un uso ridotto di pause lunghe ed esitazioni. Nei miei dati questo è evidente soprattutto per le pause, poiché le italiane producono il maggior numero di pause lunghe nel raccontare il film. La misura dei fenomeni di esitazione rileva un quadro in parte diverso, poiché il racconto di una storia presenta valori simili a quelli dei task interattivi. Anche in questo caso la differenza potrebbe essere legata alle caratteristiche del task stesso, poiché il supporto delle immagini favorisce lo sviluppo di una narrazione spezzettata per didascalie successive. In prospettiva longitudinale entrambe le misure mostrano variazione nei task monologici: tra t1 e t2 la misura legata alle pause registra un lieve aumento, mentre quella dei fenomeni di esitazione si riduce, mantenendo così il livello di fluenza globale sostanzialmente invariato. Nei task interattivi non ci sono cambiamenti sostanziali nel tempo.

Anche le apprendenti mostrano notevole variazione tra task, con tendenze diverse però rispetto a quelle delle compagne. Nel loro caso, le differenze non sono marcate per quelli monologici, ma sono piuttosto notevoli in quelli interattivi. Il task relativo alle telefonate registra le differenze maggiori rispetto agli altri tre: in questo, in cui i turni di conversazione dovrebbero alternarsi in modo rapido e contingente, le apprendenti ottengono a t1 valori piuttosto alti, con differenze più evidenti nella misura relativa alle pause. Nell'arco dei tre anni di osservazione mostrano progressi, con un graduale avvicinamento ai valori delle compagne italiane in entrambe le misure, e in modo più evidente per i task interattivi, ossia quelli in cui a t1 mostravano maggiori differenze.

Insomma l'analisi longitudinale e situazionale mostra come l'andamento delle traiettorie di sviluppo delle due misure di fluenza sia legato a due fattori: il task e l'interazione tra le due misure. Nei task interattivi le traiettorie sono piuttosto lineari, mentre in quelli monologici, dove peraltro già da t1 le

apprendenti sono in linea con la variazione nativa, seguono un andamento a zig zag. Nei task monologici, per le apprendenti come per le native, gli aumenti in una misura sono spesso bilanciati da lievi riduzioni nell'altra.

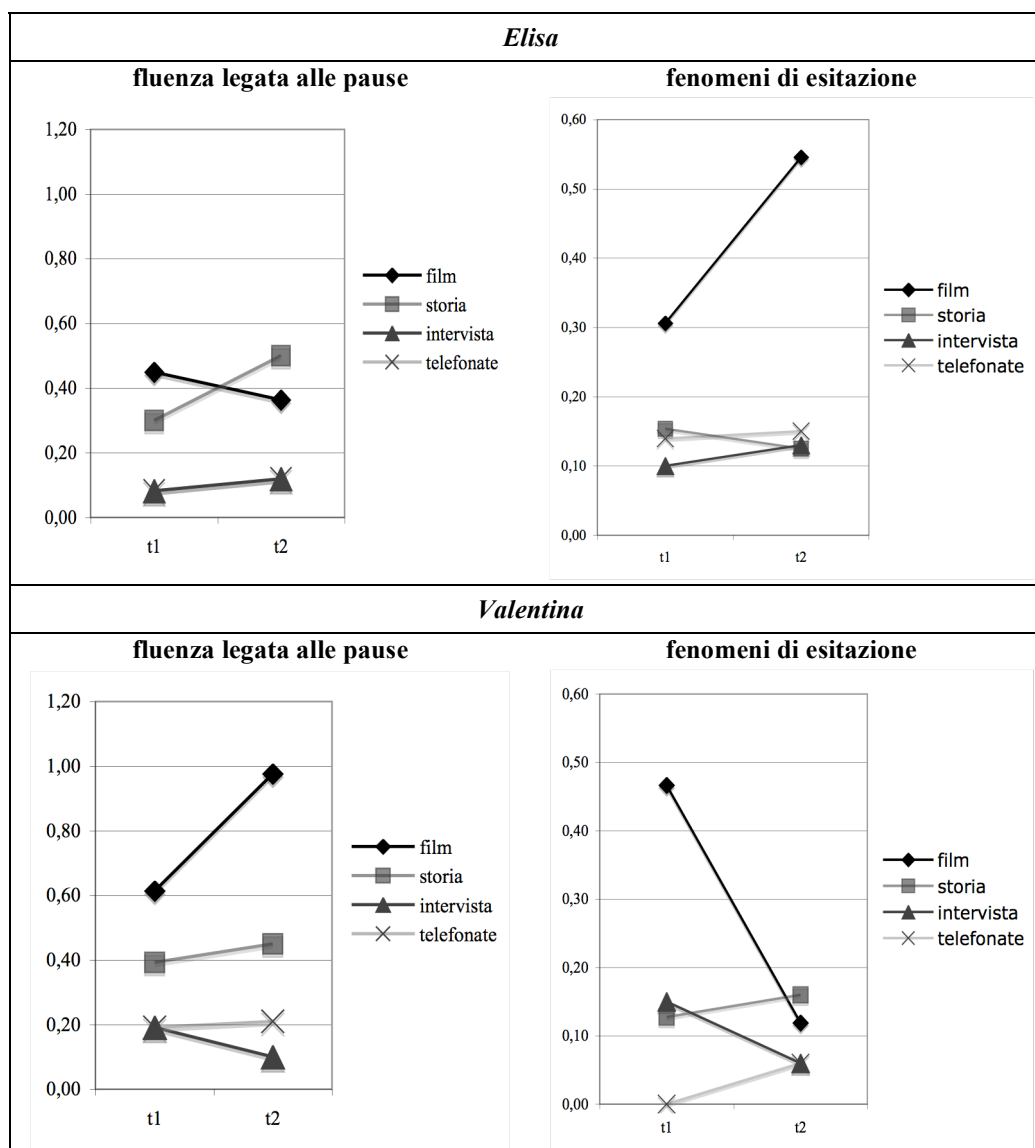
#### 4.3.3. La variazione individuale

La sezione riporta i risultati dell'analisi della variazione individuale della fluenza. Si presentano, task per task, prima i dati delle due studentesse italiane, poi quelli delle apprendenti.

Il grafico 15 riporta i percorsi individuali di ciascuna studentessa italiana. A t1 si riconferma così il quadro di variazione delineato dai valori di gruppo, e cioè che i due task monologici risultano meno fluenti dei task interattivi nella misura legata alle pause, mentre il task racconto di un film è meno fluente degli altri tre rispetto alle esitazioni. La variazione che riscontriamo tra le due ragazze è dovuta al fatto che Valentina nel raccontare il film a t2 risulta poco fluente. Abbiamo già evidenziato in § 4.1.3 come questa ragazza abbia diverse difficoltà a svolgere questo task nella seconda rilevazione. Il video le risulta poco chiaro, e pertanto la ricostruzione che ne segue è, oltre che imprecisa, estremamente faticosa, e richiede diversi interventi da parte dell'intervistatore. Questo caso mostra come la variazione della performance in un test dipenda anche da fattori esterni alla competenza linguistica o ai fattori di processing.



**Grafico 15. CAF – Fluenza – Variazione individuale – Distribuzione per task – Italiane – Pause ed esitazioni**

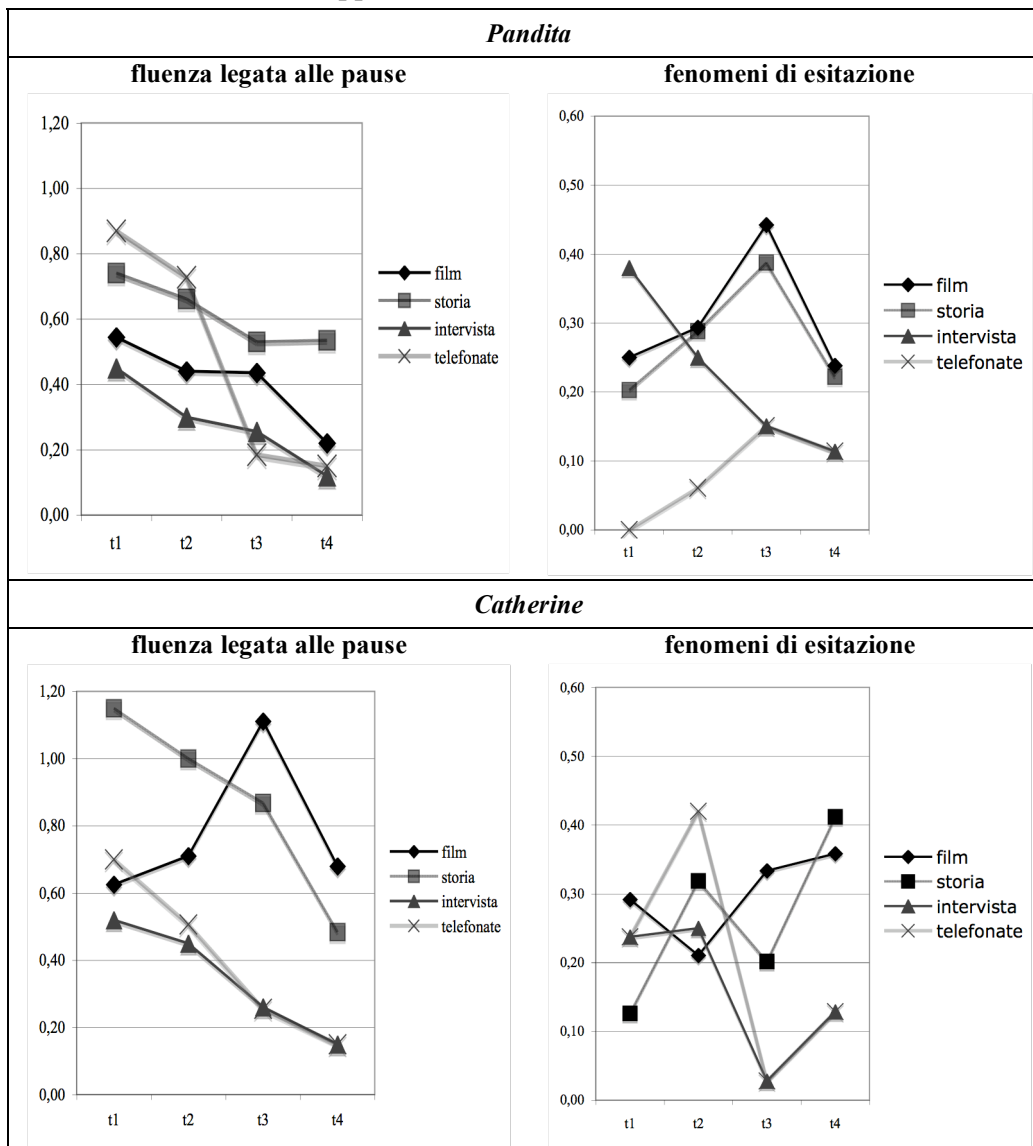


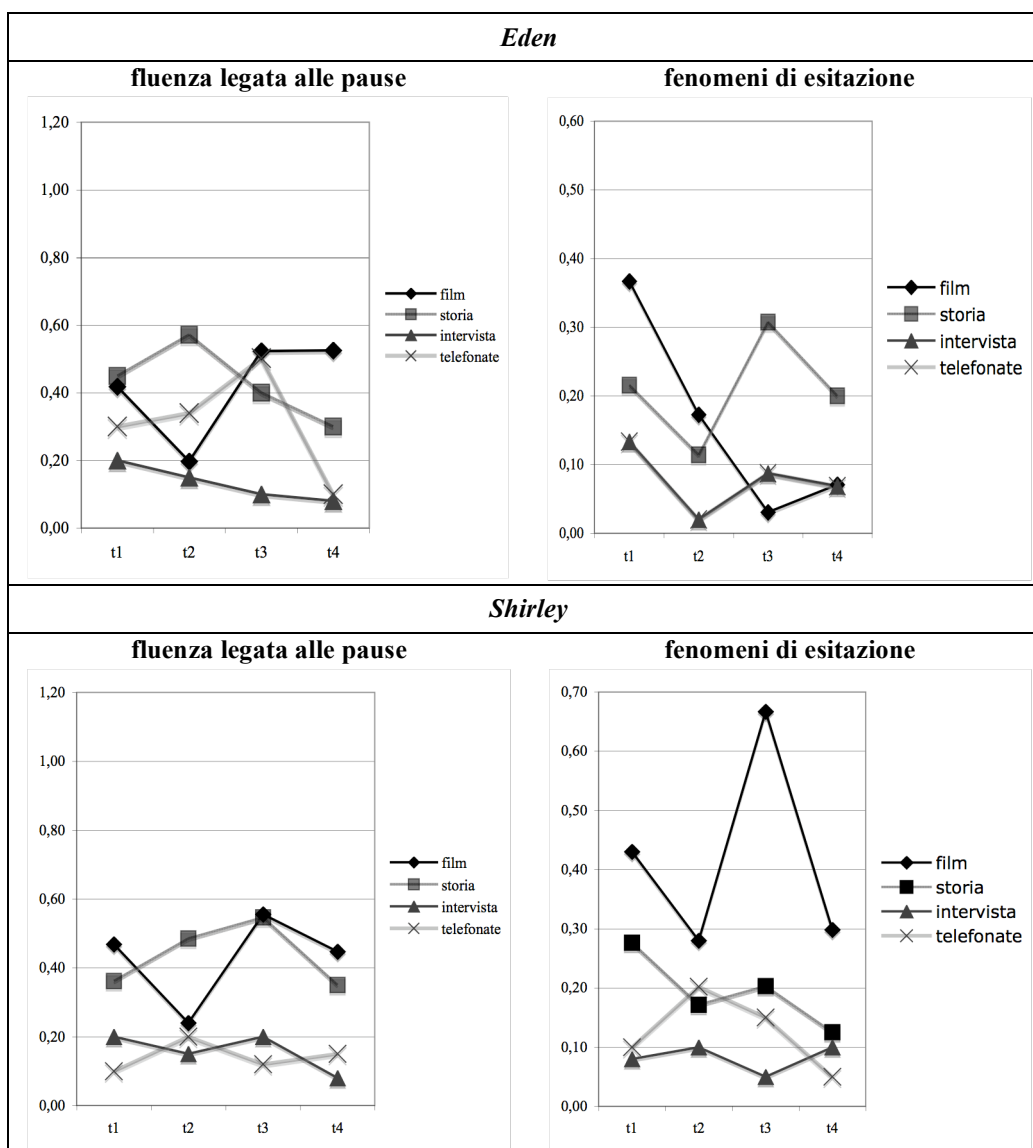
Il grafico 16 riporta i dati relativi alle quattro apprendenti. Anche per loro possiamo osservare una certa variabilità individuale. Alla prima rilevazione, la scarsa fluenza delle apprendenti rispetto alle italiane, già notata in § 4.3.1, non è uniformemente distribuita tra tutte e quattro, ma dovuta principalmente a Pandita e Catherine. Infatti queste due ragazze sono meno fluenti di Eden e Shirley in tutti i task, in particolare nelle telefonate.

Anche le traiettorie di sviluppo mostrano differenze individuali tra le apprendenti: Pandita e Catherine, meno fluenti a t1, ottengono cambiamenti più

evidenti e più lineari rispetto a Eden e Shirley. Il livello iniziale costituisce dunque un terzo elemento di variazione, che va ad aggiungersi ai due identificati in § 4.3.2, e cioè il task e le interazioni tra le due misure di fluenza.

**Grafico 16. CAF – Fluenza – Variazione individuale – Distribuzione per task – Apprendenti – Pause ed esitazioni**





Volendo adesso differenziare ulteriormente tra Pandita e Catherine, vediamo come Pandita progredisce linearmente in tutti i task, mentre Catherine è meno lineare nel racconto del film. In questo caso non si registrano sostanziali cambiamenti tra t1 e t4, e la sua traiettoria di sviluppo non è lineare poiché lievi riduzioni sono seguite da aumenti quantitativamente equivalenti. A t4 Catherine si conferma come la meno fluente del gruppo.

Poiché Eden e Shirley a t1 sono già entro il quadro di variazione delle italiane, per loro nel tempo i cambiamenti sono meno evidenti di quelli delle altre due compagne e le traiettorie di sviluppo delle misure di fluenza sono

poco lineari, senza variazioni sostanziali. Unica eccezione il task relativo alle telefonate di servizio per Eden: se a t1 si registravano alcune differenze con le italiane, nel tempo questa è l'attività in cui l'apprendente ottiene cambiamenti, allineandosi a t4 con le tendenze native.

#### 4.3.4. Riassunto

In questa sezione è stata riportata l'analisi della variazione della fluenza delle produzioni dei sei soggetti, rilevata utilizzando le due misure relative alle pause e ai fenomeni di esitazione. I dati mostrano variazione longitudinale e situazionale, sia a livello di gruppo che individuale. Vediamo con quali tendenze.

Longitudinalmente, tutte le apprendenti dimostrano cambiamenti in positivo, con un allineamento tra t1 e t4 ai valori delle compagne italiane.

Situazionalmente, le studentesse italiane mostrano differenze tra task, con quelli monologici che risultano un po' meno fluenti degli interattivi. Anche nel caso delle apprendenti c'è variazione, ma con polarità opposta. Se le differenze rispetto alle italiane sono poco notevoli per i task monologici, in quelli interattivi le apprendenti risultano meno fluenti. Nel tempo i progressi più significativi riguardano dunque senza sorprese questi ultimi task, in particolare le telefonate di servizio.

Anche individualmente c'è variazione: Pandita e Catherine, che a t1 si discostano maggiormente dalle compagne italiane rispetto a Eden e Shirley, in prospettiva longitudinale fanno i progressi maggiori e ottengono risultati più evidenti e lineari.

Infine, rispetto alle due misure considerate, i miei dati mostrano come l'andamento delle traiettorie di sviluppo della fluenza sia legato ad almeno tre fattori: il tipo di task, il livello di fluenza alla prima rilevazione, e infine il tipo di misura considerata. Gli andamenti infatti sono più lineari nei task interattivi, per le apprendenti meno avanzate, e per la misura delle pause che non nei task

monologici, per le apprendenti più avanzate e per la misura dei fenomeni di esitazione.

In diversi casi poi le due misure della fluenza si controbilanciano, con aumenti delle pause seguiti da diminuzioni nelle esitazioni. Questo dato, in linea con la letteratura (cfr. § 2.2.2.2), conferma l'ipotesi che esse rilevino diverse dimensioni di fluenza e pertanto sia utile combinarle. È anche vero però che, nei miei dati, la misura legata alle pause ottiene meno variazione individuale tra le due studentesse italiane, e per questo si può supporre sia tra le due la più precisa nella misurazione complessiva di variazione longitudinale e situazionale.

#### 4.4. *Complessità, accuratezza e fluenza CAF a confronto*

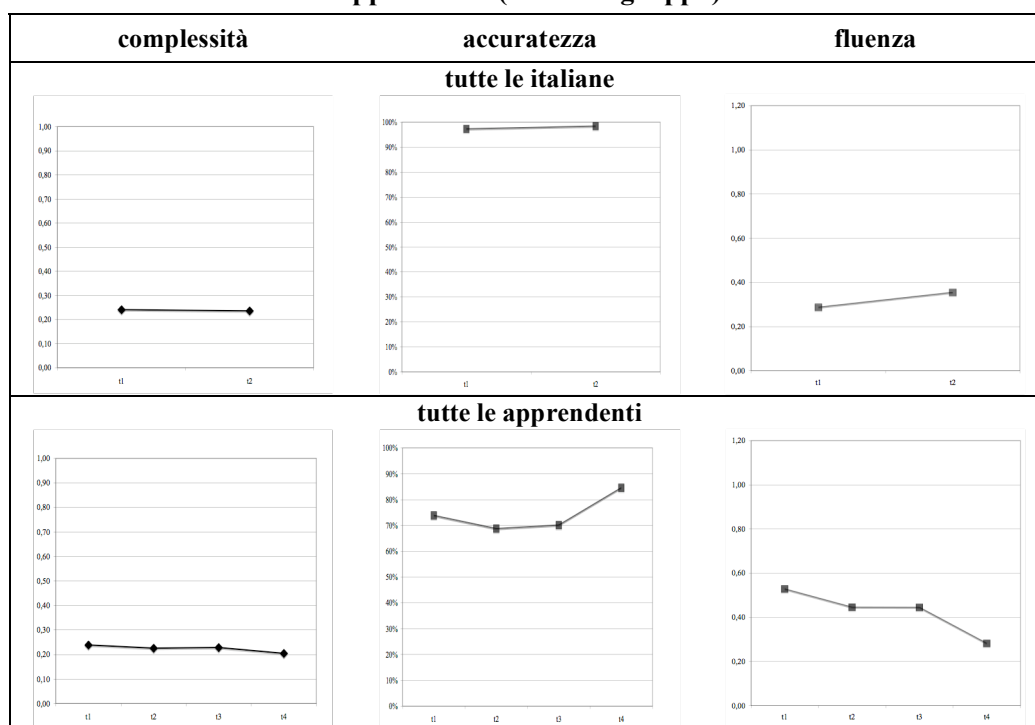
Viene qui illustrato il confronto tra gli andamenti delle tre dimensioni CAF, esaminate separatamente in §§ 4.1-4.3. Poiché, come abbiamo visto laddove sono state usate più misure alcune risultano più incisive di altre, qui ne viene selezionata una per ciascuna dimensione: la complessità sintattica viene misurata attraverso la misura subordinate per AS-unit, l'accuratezza come percentuale di AS-unit senza errori sul totale delle AS-unit prodotte, e infine la fluenza come numero medio di pause lunghe per AS-unit. Si inizierà illustrando le interazioni tra le tre dimensioni nello sviluppo longitudinale cumulativamente in tutti i task (§ 4.4.1), si continuerà con la variazione tra task (§ 4.4.2), e si concluderà con il percorso individuale di ogni soggetto (§ 4.4.3).

#### 4.4.1. La variazione longitudinale

Il grafico 17 mostra i valori medi nelle tre misure scelte per le tre dimensioni considerando i due gruppi di soggetti cumulativamente in tutti i task.

A t1 si registrano differenze tra italiane e apprendenti, in misura diversa però per ogni dimensione: le apprendenti sono complessivamente meno fluenti e accurate delle compagne italiane, mentre non ci sono differenze sostanziali in complessità. Nel corso dei tre anni, senza sorprese, accuratezza e fluenza sono le dimensioni in cui si registra il maggior progresso. A t4 le apprendenti si allineano ai valori delle compagne italiane rispetto alla fluenza, mentre rimangono alcune differenze in accuratezza.

**Grafico 17. CAF – Variazione longitudinale – Tutti i task –Italiane e apprendenti (valori di gruppo)**



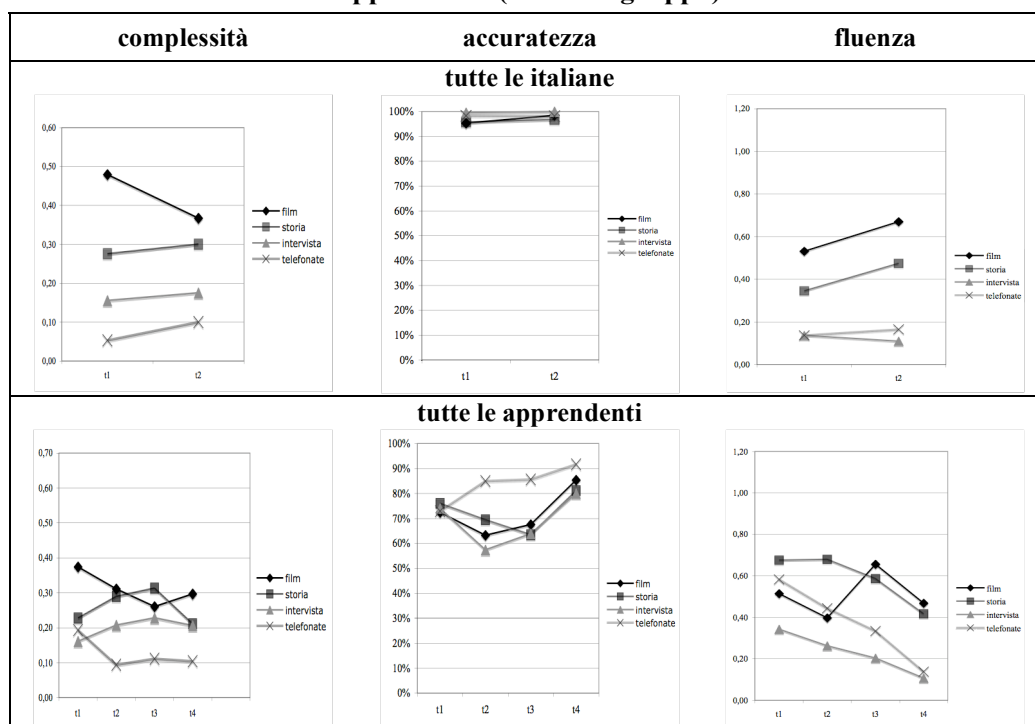
#### 4.4.2. La variazione situazionale

Il grafico 19 mostra le medie dei valori ottenuti dai due gruppi in ciascun task. I dati così disaggregati rivelano una certa variazione situazionale, arricchendo il quadro descritto in § 4.4.1.

Il gruppo di studentesse italiane dimostra variazione situazionale in tutte e tre le dimensioni, con differenze maggiori per complessità e fluenza: i task monologici risultano sintatticamente più complessi, ma un po' meno accurati e fluenti di quelli interattivi. Abbiamo già evidenziato in §§ 4.1-4.3 come tale variazione rifletta le diverse richieste dei task (cfr. § 2.2.1), con quelli monologici che favoriscono appunto la complessità e quelli interattivi la fluenza.

Nel tempo per le italiane si registra variazione: come già evidenziato dalle analisi condotte sulle singole misure (§§ 4.1-4.3) le traiettorie di sviluppo sono più o meno lineari e più o meno ascendenti in relazione al task considerato. Il confronto tra dimensioni CAF aggiunge poi un'ulteriore possibile interpretazione dell'andamento non lineare delle traiettorie di sviluppo: i possibili effetti di trade-off tra le tre dimensioni CAF. Nei task interattivi i progressi possono coinvolgere tutte e tre le dimensioni contemporaneamente, mentre nei task monologici si registrano interazioni tra le dimensioni: nel task storia un aumento della complessità è accompagnato da una riduzione della fluenza, mentre nel task film un aumento in accuratezza è accompagnato da una riduzione sia in fluenza che in complessità.

**Grafico 18. CAF – Variazione situazionale – Distribuzione per task –Italiane e apprendenti (valori di gruppo)**



Vediamo ora la variazione situazionale definita dalle apprendenti. Alla prima rilevazione le apprendenti si discostano dal quadro di variazione disegnato dalle italiane, confermando quanto osservato in §§ 4.1-4.3: non solo dimostrano meno variazione tra task, ma in alcuni casi addirittura di polarità opposta. Se nei task monologici le apprendenti alla prima rilevazione sono già piuttosto vicine ai valori delle compagne, con differenze un po' più evidenti in accuratezza, nei task interattivi, soprattutto nelle telefonate, le differenze con le italiane sono ben più evidenti, in tutte e tre le dimensioni. I task interattivi hanno livelli di complessità più alti delle italiane, allo stesso livello di quelli monologici, oltre che essere caratterizzati da una minore fluenza. Come già sottolineato nell'analisi relativa a ciascuna dimensione (§§ 4.1-4.3) i miei dati confermano dunque l'importanza della valutazione della proficiency in diverse attività e contesti, con la variabile monologico-interattivo che risulta essere particolarmente interessante.



Nel tempo le apprendenti ottengono notevoli progressi e a t4 nel complesso sono molto più vicine ai valori e al quadro di variazione definito per le italiane, con i progressi più evidenti nel task in cui mostravano più difficoltà a t1, cioè quello interattivo di apertura delle telefonate. Così come per le italiane, i dati da un lato confermano quanto mostrato nell'analisi di ogni singola dimensione, cioè che i cambiamenti nel tempo seguono traiettorie diverse sulla base del task considerato, dall'altro aggiungono un ulteriore elemento che giustifica gli andamenti più o meno lineari delle traiettorie di sviluppo: i possibili effetti di trade-off tra dimensioni CAF. Vediamone alcuni esempi.

Nei task monologici i cambiamenti nel complesso riguardano soprattutto l'accuratezza, con lievi progressi in complessità e fluenza. Nel film tra t1 e t2 complessità e accuratezza si riducono contemporaneamente, con un progresso in fluenza, tra t2 e t4, si assiste ad una lieve riduzione in complessità, accompagnata da un evidente progresso in accuratezza. La fluenza dopo l'incremento iniziale, ha un andamento non lineare senza sostanziali cambiamenti nel tempo, controbilanciando gli andamenti nelle altre due dimensioni. Diversamente, nella storia tra t1 e t3 aumentano complessità sintattica e fluenza, mentre l'accuratezza si riduce, tra t3 e t4, con un'inversione di tendenze, i resoconti risultano sintatticamente meno complessi, ma più accurati e molto più fluenti. Quest'ultima dimensione è quella che ottiene i progressi più evidenti.

Nei task interattivi invece le apprendenti si allineano ai valori delle compagne attraverso una riduzione della misura di complessità, accompagnato da un aumento in accuratezza e fluenza. Nell'intervista, senza sorprese, nel tempo i progressi più evidenti sono ottenuti in fluenza, dove i valori si allineano con quelli delle compagne italiane, con effetti di trade-off invece tra complessità e accuratezza. Se la fluenza progredisce linearmente nel tempo, complessità e accuratezza si sviluppano secondo traiettorie non lineari, spesso con polarità opposta. Infine, nel task telefonate le apprendenti realizzano i progressi più importanti, la complessità decresce allineandosi a valori più

vicini a quelli nativi, mentre accuratezza e fluenza progrediscono: le apprendenti a t4 producono meno errori e meno pause. Anche in questo caso è possibile osservare un'interazione tra complessità e accuratezza e fluenza, con una riduzione dei valori nella prima dimensione controbilanciata da progressi nelle altre due.

Ricapitolando, il quadro definito da italiane e apprendenti è chiaramente in accordo con i modelli psicolinguistici presentati in § 2.2.1.2, in cui si sostiene che le capacità di attenzione siano limitate e che pertanto nello svolgimento di un compito comunicativo, poichè il parlante decide di gestire le proprie risorse, si generano effetti di trade-off tra dimensioni. I miei dati mostrano come questo si rifletta anche nello sviluppo longitudinale: le traiettorie segnate dalle tre dimensioni variano non solo in relazione ai task ma anche ai conseguenti trade-off tra misure.

#### 4.4.3. La variazione individuale

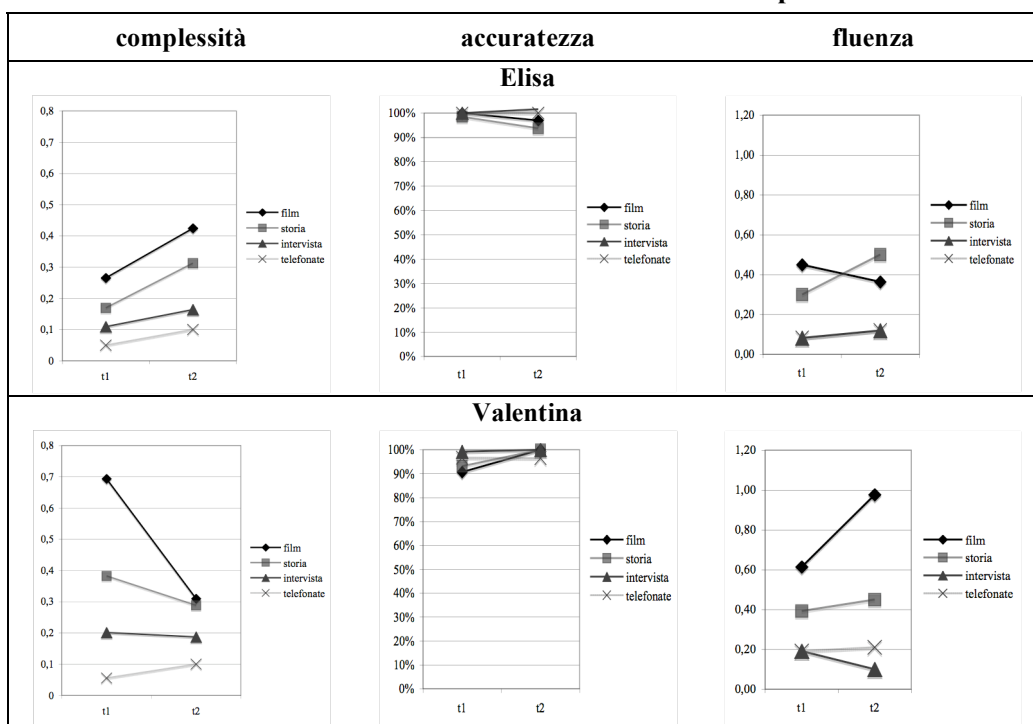
I grafici 19-20 riportano i valori ottenuti in ciascun task disaggregando i dati per soggetto. I percorsi delle singole studentesse, pur mostrando variazione individuale riflettono il quadro delineato dai valori di gruppo, con i task monologici che risultano più complessi, ma meno fluenti e un poco meno accurati di quelli interattivi. Vediamo in dettaglio i percorsi delle due studentesse italiane e delle quattro apprendenti.

Elisa e Valentina dimostrano variazione individuale nei task monologici, a t1 Elisa produce narrazioni meno complesse della compagna, risultando pertanto più fluente ed accurata. In prospettiva longitudinale le due studentesse convergono su valori simili. Da un lato Elisa complessifica la sua narrazione, con una conseguente riduzione in fluenza, Valentina, al contrario semplifica la sua narrazione, risultando anche più accurata.

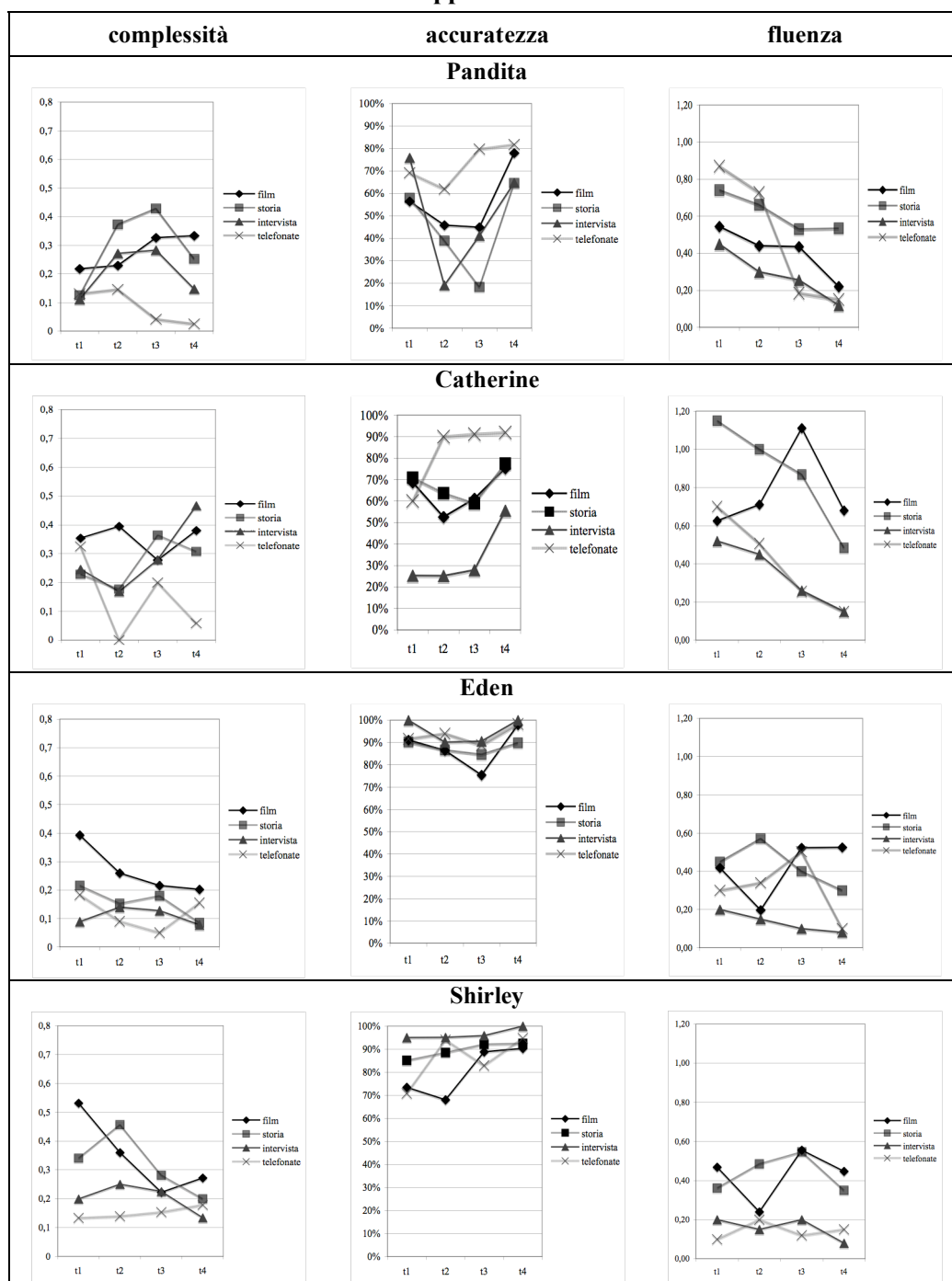
I dati mostrano come in alcuni casi il parlante possa decidere di favorire diverse dimensioni CAF nello svolgere un determinato compito. Tale scelta

può essere legata, come nel caso appena descritto, a scelte individuali rispetto alla strategia narrativa: poiché le capacità di attenzione sono limitate, se il parlante sceglie di realizzare un racconto complesso, ci saranno riduzioni in fluenza e accuratezza, come nel caso di Valentina, se sceglie invece di realizzare un racconto più semplice, la riduzione in complessità sarà accompagnata da un aumento di fluenza e di accuratezza, come nel caso di Elisa. In altri casi ancora, come nel task film a t2 per Valentina, la variazione può essere legata a fattori esterni al task: abbiamo già visto in § 4.1-4.3 come la studentessa incontri diverse difficoltà nello svolgere questo compito, di conseguenza la sua narrazione è semplice, frammentaria, caratterizzata da pause ed esitazioni.

**Grafico 19. CAF – Variazione individuale – Distribuzione per task – Italiane**



**Grafico 20. CAF – Variazione individuale – Distribuzione per task – Apprendenti**



Il grafico 20, che riporta i risultati relativi alle quattro apprendenti, mostra variazione individuale, con Pandita e Catherine che a t1 dimostrano meno variabilità tra task rispetto a Eden e Shirley, risultando nel complesso meno avanzate delle compagne. Per tutte le apprendenti comunque le differenze più evidenti riguardano i task interattivi, in particolare quello relativo alle telefonate. Nel tempo tutte le ragazze realizzano diversi progressi, anche se variano da individuo a individuo dimensioni coinvolte e andamenti delle traiettorie di sviluppo. Vediamo nel dettaglio i percorsi individuali delle quattro apprendenti. Si illustreranno task per task prima i dati relativi a Pandita e Catherine, poi quelli di Eden e Shirley.

Pandita e Catherine, le meno avanzate a t1, dimostrano i progressi più significativi, con un sostanziale allineamento ai valori nativi a t4. Nel tempo c'è più variabilità tra task e la direzione di tale variazione è in sintonia con le produzioni delle studentesse italiane.

Nei task monologici per Pandita la crescita in complessità sintattica e l'acquisizione di strutture linguistiche generano un effetto di trade-off con l'accuratezza, mentre la fluenza cresce linearmente. Analogo il percorso di sviluppo di Catherine, anche se meno lineare. Nel suo caso possiamo osservare trade-off tra tutte e tre le dimensioni. Anche per lei una riduzione in accuratezza è il risultato di uno sviluppo interlinguistico, diversamente da Pandita è però accompagnato da cambiamenti poco lineari anche in fluenza.

Nel task intervista Pandita e Catherine seguono invece percorsi di sviluppo fra loro diversi. Per Pandita non ci sono sostanziali differenze rispetto agli andamenti segnati dai task monologici: in una prima fase l'apprendente cresce soprattutto in complessità e fluenza, con un trade-off rispetto all'accuratezza, in una seconda fase, riavvicinandosi ai valori nativi, decresce in complessità, avendo imparato a diluire meglio le unità informative e mantenere lo scambio più fluido, mentre aumentano accuratezza e fluenza. Catherine in questo task segue al contrario un andamento più lineare, segnando un graduale progresso in tutte e tre le dimensioni.

Nel task relativo alle telefonate, sia Pandita che Catherine a t1 si discostano notevolmente dalle italiane, in prospettiva longitudinale i loro percorsi di sviluppo si differenziano. Pandita progredisce in tutte e tre le dimensioni seguendo traiettorie lineari: riduce la complessità sintattica e progredisce in accuratezza e fluenza. In questo task l'effetto di trade-off è legato a un avvicinamento anche pragmatico alle modalità di apertura delle telefonate tipiche dei nativi. Per Catherine invece le traiettorie di sviluppo CAF non sono lineari, nel suo caso i progressi in complessità e accuratezza generano un trade-off con la fluenza.

Le traiettorie di sviluppo di Eden e Shirley, relativamente più avanzate già a t1, hanno un andamento diverso da quello delle due compagne. Vediamone le tendenze.

Nei task monologici, si assiste ad una riduzione della complessità sintattica in termini di subordinazione: l'analisi qualitativa presentata in § 4.1.1.1 mostra però come tale riduzione sia il risultato, nel caso di Eden dell'uso di clausole più lunghe, con un progresso in complessità in termini di lunghezza, in quello di Shirley della scelta di uno stile narrativo più breve e semplice. Le diverse cause di riduzione dei valori di complessità in termini di subordinazione è controbilanciata da diversi andamenti della misura di accuratezza: nel caso di Eden la complessificazione in termini di lunghezza è accompagnata tra t1 e t4 da una parziale riduzione in accuratezza, nel caso di Shirley, la riduzione della complessità sintattica come conseguenza di una strategia narrativa più semplice porta ad un aumento dell'accuratezza. Tale sviluppo longitudinale genera anche nella fluenza un andamento non lineare e piuttosto simile per le due apprendenti, indicando a t4 a un lieve progresso rispetto alla prima rilevazione, già in linea comunque con i valori delle compagne italiane.

Nel task intervista sia Eden che Shirley seguono traiettorie di sviluppo lineari in tutte le dimensioni, senza effetti di trade-off.

Nel task relativo alle telefonate invece le due apprendenti seguono percorsi meno lineari, oltre che diversi tra loro. Eden, così come Catherine,

riduce la complessità sintattica con un lieve effetto di trade-off rispetto alla fluenza. Shirley segue invece traiettorie di sviluppo non lineari in tutte e tre le dimensioni, con la complessità che registra un lieve aumento.

I dati fin qui discussi ci permettono di riconfermare alcune osservazioni relative agli andamenti delle misure CAF e alle interazioni che si realizzano durante la realizzazione di un determinato task o come conseguenza dello sviluppo longitudinale. È chiaro che, come nel caso delle nostre apprendenti, la traiettoria di sviluppo di queste dimensioni può procedere in modo non lineare sulla base di almeno quattro elementi: livello iniziale, task considerato, variazione individuale e trade-off tra dimensioni. Gli andamenti non lineari nelle traiettorie di sviluppo e le interazioni tra dimensioni suggeriscono come la combinazione delle misure CAF permetta una descrizione più affidabile del profilo di proficiency di un apprendente in un determinato task.

#### 4.4.4. Riassunto

In questa sezione è stata riportata l'analisi delle interazioni tra le dimensioni CAF. Per la complessità sintattica è stata considerata la misura subordinate per AS-unit, per l'accuratezza la percentuale di AS-unit senza errori sul totale delle AS-unit e infine per la fluenza il numero di pause lunghe per AS-unit.

Longitudinalmente, per tutte e quattro le apprendenti i dati dimostrano progresso, nel senso di un avvicinamento alle italiane in tutte e tre le dimensioni, con cambiamenti più evidenti in accuratezza e fluenza, cioè le dimensioni che dimostravano le maggiori differenze a t1.

Situazionalmente, c'è variazione tra task. Per le studentesse italiane c'è variazione in tutte e tre le dimensioni: i task monologici risultano sintatticamente più complessi, ma un po' meno accurati e fluenti dei task interattivi. Anche per le apprendenti c'è un po' di variazione tra task a t1, ma le differenze sono meno evidenti che per le italiane, in diversi casi con una polarità inversa: ad esempio, rispetto a complessità e fluenza spesso le

apprendenti risultano più complesse e meno fluenti nei task interattivi che in quelli monologici. Nel tempo tutte le apprendenti realizzano progressi, più evidenti nei task interattivi e rispetto a accuratezza e fluenza, e a t4 sono più vicine al quadro di variazione disegnato dalle compagne italiane. L'andamento non lineare delle misure CAF è legato in alcuni casi a interazioni tra le dimensioni stesse: ad esempio l'accuratezza e la complessità nei task monologici, all'aumentare dell'una si riduce l'altra.

Anche individualmente vi sono differenze quantitative e qualitative nelle traiettorie di sviluppo percorse da ciascuna soggetto. Per le italiane, nei task monologici la scelta personale di uno stile narrativo più complesso, informativamente denso e coeso, piuttosto che uno stile più conciso e veloce, genera un trade-off tra complessità, accuratezza e fluenza, con la narrazione più complessa che risulta un po' meno accurata e fluente. Anche per le apprendenti c'è variazione individuale, con un alcune differenze però rispetto alle italiane. I dati individuali mostrano infatti come la variazione dipenda, oltre che da task e interazioni tra dimensioni, anche dal livello iniziale. Questo può determinare infatti andamenti più o meno lineari nelle misure, oltre che favorire un tipo di cambiamento piuttosto che un altro.

Infine, rispetto alle misure considerate, gli andamenti non lineari nelle traiettorie di sviluppo e le interazioni tra dimensioni suggeriscono come queste misure non possano essere considerate indicatori diretti della competenza, poiché un aumento o una diminuzione è un cambiamento in positivo e in negativo sulla base del livello di partenza e del target nativo, né possono essere usate singolarmente, poiché è la loro interazione che permette la descrizione affidabile del profilo di proficiency di un apprendente in un determinato task.



#### 4.5. *Le valutazioni QCER*

La sezione riporta i risultati delle valutazioni espresse dai due valutatori sulla base di tre descrittori del QCER: coerenza-coesione, correttezza grammaticale e fluenza. Sono qui considerati i dati individuali di tre task del corpus (storia, intervista e apertura delle telefonate di servizio) prodotti alla prima e all'ultima rilevazione (t1 e t2 per le italiane, t1 e t4 per le apprendenti). Si illustreranno prima i risultati della variazione longitudinale per ciascun descrittore (§ 4.5.1), quindi si presenteranno alcune osservazioni sulla variazione situazionale (§ 4.5.2).

##### 4.5.1. La variazione longitudinale

La tabella 24 mostra le valutazioni ottenute dalle due studentesse italiane. Elisa e Valentina non ottengono sempre il punteggio massimo, in particolare per il descrittore coerenza-coesione, dove si collocano in fascia B. Longitudinalmente registrano alcuni progressi, con differenze individuali rispetto ai descrittori coinvolti: Elisa mostra alcuni cambiamenti in termini di accuratezza e Valentina soprattutto in termini di fluenza.

**Tabella 24. QCER – Variazione longitudinale – Italiane**

	Elisa		Valentina	
	t1	t2	t1	t2
<b>coerenza-coesione</b>				
<b>storia</b>	C1	C1	B2+	B2
<b>intervista</b>	B2+	B2+	B2+	C1
<b>telefonate</b>	C2	C1	C1	C1
<b>accuratezza</b>				
<b>storia</b>	C2	C2	C2	C2
<b>intervista</b>	C1	C2	C1	C2
<b>telefonate</b>	C2	C2	C2	C2
<b>fluenza</b>				
<b>storia</b>	C2	C2	C1	B2
<b>intervista</b>	C2	C1	C1	C2
<b>telefonate</b>	C2	C2	C1	C2

La tabella 25 riporta i dati relativi alle apprendenti. Come atteso ricevono valutazioni più basse rispetto alle compagne italiane, Pandita e Catherine risultano meno avanzate in tutti e tre i descrittori e si collocano al livello B1 o B1+, Eden e Shirley sono più avanzate e si collocano tra i livelli B2 e C2 in tutti i descrittori, con le valutazioni migliori in accuratezza. In prospettiva longitudinale tutte le apprendenti registrano progressi, anche se con una certa variazione individuale rispetto all'intensità e ai descrittori considerati.

Pandita è l'apprendente che ottiene i cambiamenti maggiori in tutte le dimensioni, soprattutto in fluenza, dove si assiste a un passaggio dal livello B al livello C. Catherine mostra meno variazione longitudinale della compagna, con i cambiamenti più evidenti in termini di accuratezza. Eden e Shirley progrediscono in tutte le dimensioni, con i cambiamenti maggiori in accuratezza per Eden e in coerenza-coesione e fluenza per Shirley.

Rispetto ai descrittori selezionati, si sono evidenziati alcuni limiti nell'uso del QCER. I valutatori infatti hanno avuto difficoltà a stabilire una valutazione per il descrittore relativo a coerenza-coesione: le discussioni per un

accordo erano più lunghe e in tre casi non è stato espresso alcun giudizio. Vedremo in § 4.5.2 come le difficoltà maggiori riguardino i task interattivi.

**Tabella 25. QCER – Variazione longitudinale – Apprendenti**

	Pandita		Catherine		Eden		Shirley	
	t1	t4	t1	t4	t1	t4	t1	t4
<i>coerenza-coesione</i>								
<b>storia</b>	B1	C1	B1	B1	B2	C1	B2+	C1
<b>intervista</b>	B2	B2+	B1+	C1	?	C1	B2+	C1
<b>telefonate</b>	B1	B2+	?	B2	?	B2+	B2+	C1
<i>accuratezza</i>								
<b>storia</b>	B1+	B2+	B1+	B1+	B2+	C1	C1	C1
<b>intervista</b>	B2	B2	B1+	B2+	C1	C2	C1	C2
<b>telefonate</b>	B2	C1	B2	C1	C1	C1	C1	C1
<i>fluenza</i>								
<b>storia</b>	B1+	C1	B1	B1	B1+	B2+	C1	C2
<b>intervista</b>	B2+	C1	B1+	C1	C1	C1	B2+	C2
<b>telefonate</b>	B1+	C1	B1+	B1+	C1	C2	C1	C2

#### 4.5.2. La variazione situazionale

Per studentesse italiane e apprendenti a t1 c'è variazione tra task, con i task interattivi che ottengono valutazioni in genere migliori. In prospettiva longitudinale i progressi si differenziano nei diversi task e coinvolgono diverse dimensioni, con una certa variazione individuale.

La tabella 26 riporta le valutazioni espresse per le studentesse italiane. Elisa e Valentina ottengono le valutazioni migliori nel task interattivo di apertura delle telefonate, mentre nei task storia e intervista il descrittore coerenza-coesione registra valutazioni più basse. Il task di apertura delle telefonate di servizio ottiene le valutazioni migliori in tutte e tre le dimensioni considerate.

**Tabella 26. QCER – Variazione situazionale – Italiane**

	Elisa		Valentina	
	t1	t2	t1	t2
	<i>storia</i>			
<b>coerenza-coesione</b>	C1	C1	B2+	B2
<b>accuratezza</b>	C2	C2	C2	C2
<b>fluenza</b>	C2	C2	C1	B2
	<i>intervista</i>			
<b>coerenza-coesione</b>	B2+	B2+	B2+	C1
<b>accuratezza</b>	C1	C2	C1	C2
<b>fluenza</b>	C2	C1	C1	C2
	<i>telefonate</i>			
<b>coerenza-coesione</b>	C2	C1	C1	C1
<b>accuratezza</b>	C2	C2	C2	C2
<b>fluenza</b>	C2	C2	C1	C2

La tabella 27 riporta le valutazioni per le quattro apprendenti. Come per le italiane, possiamo osservare variazione situazionale, con i task interattivi che ottengono valutazioni in genere più alte del task monologico.

A t1 Pandita raggiunge le valutazioni migliori per l'intervista in tutti i descrittori e per le telefonate nel descrittore accuratezza (B2). Così come per le italiane, la storia è il task che ottiene le valutazioni più basse (B1 e B1+). In prospettiva longitudinale, l'apprendente mostra un generale miglioramento, il più significativo riguarda le telefonate, dove per accuratezza e fluenza passa da B2 e B1+ a C1.

Catherine invece non mostra variazione situazionale a t1, collocandosi tra B1 e B1+ in tutti i task e per tutti i descrittori. In prospettiva longitudinale dimostra progressi soprattutto nel task intervista, mentre in quello relativo alle telefonate ci sono alcuni cambiamenti in termini di accuratezza. Non ci sono progressi nel task monologico.

A t1 Eden dimostra una certa variazione tra task, con la storia che ottiene valutazioni nella fascia B e i task interattivi nella fascia C. In prospettiva

longitudinale, i progressi riguardano soprattutto il task storia, mentre nell'intervista progredisce solo l'accuratezza, e nelle telefonate la fluenza.

Shirley è la studentessa che a t1 ottiene le valutazioni migliori, con valutazioni un po' più basse nel task intervista e per il descrittore coerenza-coesione. In prospettiva longitudinale l'apprendente migliora in tutti i task, senza sorprese, soprattutto nell'intervista e per i i descrittori coerenza-coesione e fluenza.

**Tabella 27. QCER – Variazione situazionale – Apprendenti**

	<b>Pandita</b>		<b>Catherine</b>		<b>Eden</b>		<b>Shirley</b>	
	<b>t1</b>	<b>t4</b>	<b>t1</b>	<b>t4</b>	<b>t1</b>	<b>t4</b>	<b>t1</b>	<b>t4</b>
	<i>storia</i>							
<b>coerenza-coesione</b>	B1	C1	B1	B1	B2	C1	B2+	C1
<b>accuratezza</b>	B1+	B2+	B1+	B1+	B2+	C1	C1	C1
<b>fluenza</b>	B1+	C1	B1	B1	B1+	B2+	C1	C2
	<i>intervista</i>							
<b>coerenza-coesione</b>	B2	B2+	B1+	C1	?	C1	B2+	C1
<b>accuratezza</b>	B2	B2	B1+	B2+	C1	C2	C1	C2
<b>fluenza</b>	B2+	C1	B1+	C1	C1	C1	B2+	C2
	<i>telefonate</i>							
<b>coerenza-coesione</b>	B1	B2+	?	B2	?	B2+	B2+	C1
<b>accuratezza</b>	B2	C1	B2	C1	C1	C1	C1	C1
<b>fluenza</b>	B1+	C1	B1+	B1+	C1	C2	C1	C2

Così come anticipato in §4.5.1 il QCER mostra però alcuni limiti nella valutazione della coerenza-coesione nei task interattivi. In tre casi i valutatori non riescono a raggiungere un accordo. In pratica, i valutatori hanno la sensazione che l'apprendente si trovi a un certo livello, ma le definizioni del descrittore non siano sempre adatte a descrivere la qualità della performance. Nel nostro caso infatti la descrizione della coerenza-coesione è formulata in termini di complessificazione e non è pertanto adatta a descrivere le richieste di un compito altamente interattivo. I miei dati suggeriscono dunque il bisogno di specificare la formulazione di questo descrittore in relazione al task

considerato, monologico o interattivo. Limiti di questo tipo per l'applicazione del QCER al testing sono già stati evidenziati dalla letteratura (*linking problem*, cfr. Alderson *et al.*, 2004; Weir, 2005; § 2.3.2.2; § 3.5.2).

#### 4.5.3. Riassunto

Riguardo alla descrizione delle valutazioni espresse da due valutatori a partire dai descrittori QCER selezionati (coerenza-coesione, accuratezza e fluenza) su tre task (storia, intervista e telefonate) realizzati alla prima e all'ultima rilevazione, i dati mostrano variazione.

Longitudinalmente, tutte le apprendenti realizzano progressi. A t1 le apprendenti, come atteso, ottengono valutazioni inferiori rispetto a quelle delle compagne italiane. Nel tempo accuratezza e fluenza ottengono i cambiamenti più significativi.

Situazionalmente, si registrano a t1 per tutti i soggetti differenze tra task, con quello monologico che riceve le valutazioni più basse. Interessante notare che le studentesse italiane non raggiungono sempre il livello C. Nel tempo le apprendenti non progrediscono allo stesso ritmo in tutti i task, i cambiamenti più significativi riguardano infatti quelli interattivi.

Anche individualmente c'è variazione. A t1 Pandita e Catherine si collocano nella fascia B e Eden e Shirley più vicine alla fascia C. Nel tempo poi i progressi si differenziano da apprendente a apprendente per intensità e per task e descrittori coinvolti.

Infine, per quel che riguarda i descrittori QCER, quelli selezionati risultano nel complesso piuttosto efficaci, anche se per coerenza-coesione non sono sempre sensibili alla variazione tra task. I valutatori hanno difficoltà nell'esprimere la loro valutazione su questa dimensione nei task interattivi, in modo più evidente per le telefonate. Il descrittore QCER risulta poco sensibile alle caratteristiche di questa attività. Sarebbe dunque opportuno per chi utilizza il QCER ai fini della valutazione avere informazioni più dettagliate sulla

qualità della produzione attesa per ogni livello differenziando per task, con la variabile monologico-interattivo che si rivela particolarmente interessante.

#### 4.6. *Misure CAF e valutazioni QCER a confronto*

La sezione mette a confronto i risultati dell'analisi della variazione di complessità, accuratezza e fluency dei sei soggetti, ottenuti dall'applicazione delle misure CAF e dalle valutazioni soggettive QCER, prima esaminate separatamente (§§ 4.1-4.4 e § 4.5). Sono qui considerati i dati individuali di tre task del corpus, racconto di una storia ad immagini, intervista e apertura delle telefonate di servizio, realizzati a t1 e t2 dalle studentesse italiane e a t1 e t4 dalle apprendenti. Si discuterà prima la variazione longitudinale (§ 4.6.1), poi la variazione situazionale (§ 4.6.2). L'obiettivo è indagare le relazioni che intercorrono tra i due tipi di misurazioni e rilevare somiglianze, differenze ed eventuali punti di complementarità.

##### 4.6.1. *La variazione longitudinale*

Nel complesso, i quadri di variazione longitudinale definiti dai due tipi di misurazione sono coerenti. Le studentesse italiane, senza sorprese, sono più avanzate delle apprendenti, sia CAF che QCER mostrano poi come non sempre raggiungono il punteggio massimo e come ci sia variazione individuale. In prospettiva longitudinale entrambi i tipi di misurazione rilevano lievi cambiamenti nel tempo. Anche per le apprendenti CAF e QCER disegnano quadri nel complesso coerenti. A t1 entrambe le misurazioni mostrano come Pandita e Catherine siano meno avanzate di Eden e Shirley in tutte le dimensioni e come, nel tempo, tutte le ragazze registrano progressi.

Ciononostante, nell'analisi della variazione longitudinale il confronto tra CAF e QCER mette in luce alcune differenze. La prima riguarda l'ampiezza

dello sviluppo delle singole dimensioni: le misure quantitative CAF mostrano come le differenze principali tra studentesse italiane e apprendenti riguardino accuratezza e fluenza, mentre QCER a t1 non rileva questa differenza. In prospettiva longitudinale però i due strumenti concordano maggiormente, così come CAF anche QCER registra i progressi più evidenti in fluenza. Questa prima differenza costituisce un punto di reciproca complementarità: l'analisi CAF infatti permette di identificare quali dimensioni ottengano il maggior cambiamento a livello avanzato, offrendo spunti utili per una miglior calibratura dei descrittori QCER.

La seconda differenza riguarda invece la direzione e l'andamento dello sviluppo delle dimensioni. Le traiettorie di sviluppo CAF non sono sempre lineari, abbiamo visto in §§ 4.1-4.4 come questo sia il risultato dell'interazione di una serie di elementi, livello interlinguistico, sviluppo, task considerato e interazione tra dimensioni. Queste misure pertanto non sono indicatori diretti della competenza, ma devono essere interpretate a partire dalle reciproche relazioni e dal confronto con parlanti nativi. Le valutazioni QCER confermano l'idea che il progresso non sia indiscriminatamente un incremento in complessità o accuratezza: in tutti questi casi infatti segnano un progresso, guidando dunque l'interpretazione di CAF e confermando le osservazioni fatte nell'analisi delle singole misure. Approfondiremo meglio questa seconda differenza in § 4.6.2 descrivendola in rapporto ai singoli task.

#### 4.6.2. La variazione situazionale

Il confronto tra misure CAF e valutazioni QCER nella descrizione della variazione situazionale evidenzia come i due tipi di misurazione, nonostante descrivano un quadro nel complesso coerente, mostrano alcune differenze.

Innanzitutto il QCER si dimostra poco sensibile alla variazione situazionale. I risultati dell'analisi CAF mostrano come in genere le apprendenti siano più distanti dal comportamento nativo nel task interattivo di



apertura delle telefonate, e più vicine alle compagne italiane nell'intervista e nel task monologico. Il QCER non riflette esattamente lo stesso quadro, le apprendenti ottengono valutazioni equivalenti in tutti i task e spesso migliori in quelli interattivi. Anche longitudinalmente, se CAF mostra maggior progresso in alcuni task e dimensioni, QCER rileva un miglioramento complessivo in tutti i task e tutte le dimensioni. In questo senso è evidente come QCER e CAF differiscono nel registrare le variazioni tra task. Vediamo più in dettaglio le differenze che emergono dal confronto tra CAF e QCER nei singoli task.

Nei task monologici le differenze riguardano gli andamenti di complessità e accuratezza. Nel caso dell'accuratezza, per Pandita, Catherine e Eden la misura CAF segue traiettorie ad U, mentre per Shirley lo sviluppo è lineare. Abbiamo visto in § 4.2 come questo sia legato all'interazione tra misura scelta e sviluppo interlinguistico. Il valutatore, al contrario, più attento al tipo di errore, piuttosto che al numero, probabilmente combina in qualche modo nella sua interpretazione i costrutti sviluppo e accuratezza, registrando così progressi in accuratezza per tutte e quattro le apprendenti. In questo caso dunque QCER e CAF si integrano e possono guidare l'uno l'interpretazione dell'altro. Rispetto alla complessità invece, abbiamo osservato in § 4.1.3 come le traiettorie di sviluppo CAF delle apprendenti si differenziano sulla base del livello di partenza iniziale, per Pandita e Catherine, le meno avanzate, i progressi sono in termini di subordinazione, mentre per Eden e Shirley in lunghezza o dovuti alla sperimentazione di una più ampia varietà di stili discorsivi. Abbiamo già visto come questo dato conferma l'Ipotesi di Predizione dello Sviluppo, secondo la quale a livelli più avanzati la complessificazione non riguarda tanto la quantità di subordinate usate, ma piuttosto la varietà o l'uso di strutture più complesse a livello di clausola. QCER segna progressi per tutte le ragazze, mostrandosi dunque sensibile ad entrambi i tipi di complessificazione. Anche in questo caso dunque QCER e CAF integrano l'uno l'interpretazione dell'altro.

Anche nei task interattivi le differenze riguardano la dimensione complessità sintattica. In questi task i valutatori hanno avuto alcune difficoltà

nell'applicazione del descrittore QCER coerenza-coesione, in tre casi addirittura non sono stati in grado di arrivare ad un accordo. Questo sembra legato a una percepita discrepanza tra la descrizione del QCER e le richieste del task interattivo. L'analisi CAF mostra infatti come questo task favorisca la produzione di AS-unit semplici con una bassa densità informativa, il descrittore QCER al contrario definisce la coerenza-coesione in termini di complessificazione, risultando pertanto poco adatto alla valutazione del task telefonate. In questo caso i profili di variazione tra italiane e apprendenti definiti dalle misure CAF potrebbe essere un utile punto di partenza per migliorare la formulazione del descrittore QCER coerenza-coesione.

#### 4.6.3. Riassunto

Messe a confronto le analisi CAF e le valutazioni QCER, entrambe rilevano variazione nelle produzioni di studentesse italiane e apprendenti. A volte i due tipi di misurazione rilevano le stesse tendenze, altre i due approcci risultano complementari, con uno che guida l'interpretazione dell'altro. Il QCER complementa CAF nell'interpretazione dell'andamento delle misure di complessità sintattica e di accuratezza in quelli monologici. Sono questi i casi, ad esempio, dell'andamento non lineare delle misure CAF di complessità sintattica nei task interattivi o di accuratezza in quelli monologici. CAF invece complementa QCER nella descrizione della variazione situazionale. Sono questi i casi del descrittore coerenza-coesione nel task telefonate e delle indicazioni che CAF può dare rispetto alle dimensioni che variano a livello avanzato in relazione ai task.



## **CAPITOLO 5**

### **CONCLUSIONE**

Questa ricerca è stata condotta con lo scopo di osservare da un lato la variabilità longitudinale e situazionale dell'interlingua di apprendenti avanzati impegnati nello svolgimento di quattro task comunicativi, e dall'altro di verificare la relazione che intercorre tra due tipi di misurazione: quella analitica che rileva la complessità, l'accuratezza e la fluenza del parlato (CAF), impiegata nei cosiddetti *developmental index studies* di stampo acquisizionale, e quella più soggettiva basata sui descrittori del *Quadro Comune Europeo di Riferimento* (QCER) maggiormente legata all'ambito del testing.

Il corpus analizzato nella tesi è costituito dalle produzioni di quattro studentesse immigrate di diversa nazionalità, raccolte in quattro rilevazioni nell'arco di tre anni, integrato con due rilevazioni di due parlanti native con simili caratteristiche, che costituiscono la pietra di paragone per il confronto con le apprendenti.

Nel presente capitolo si tratterà un quadro sintetico dei risultati ottenuti dall'analisi, nel tentativo di rispondere alle domande che hanno guidato la

ricerca. Si proverà infine a individuare spunti di applicazione dei risultati agli ambiti della ricerca acquisizionale, del testing e della didattica.

### *5.1. Riassunto*

Rivediamo dunque le domande alla base della ricerca (cfr. § 1.3), per poi procedere alla presentazione delle risposte ricavate dall'analisi dei dati.

#### **Come variano complessità, accuratezza e fluenza nel corso del tempo mano a mano che progredisce l'interlingua?**

Alla prima rilevazione, senza sorprese, l'analisi cumulativa su tutte le produzioni rileva differenze significative tra i due gruppi delle studentesse italiane e delle apprendenti, in misura diversa però per ogni dimensione: se in complessità non sono sostanziali, in fluenza ed accuratezza sono piuttosto evidenti, poiché le apprendenti ottengono valori inferiori a quelli delle compagne italiane. Nel tempo il gruppo delle apprendenti registra variazione longitudinale in tutte e tre le dimensioni, con progresso più evidente nell'accuratezza e nella fluenza, rispetto alla complessità. All'ultima rilevazione le ragazze sono perciò più vicine ai valori ottenuti dalle compagne, nonostante permangano alcune differenze in accuratezza.

Le traiettorie di sviluppo di CAF non sono però lineari per tutte le dimensioni. Nel caso della complessità e dell'accuratezza i dati di gruppo rivelano come il livello di competenza linguistica avanzato delle quattro apprendenti possa costituire una prima spiegazione degli andamenti più o meno lineari. Nel caso della complessificazione sintattica, in accordo con quanto suggerito dall'Ipotesi di Predizione dello Sviluppo proposta per esempio da Cooper (1976), Monroe (1975), Sharma (1980) e più recentemente Wolfe-Quintero *et al.* (1998), le apprendenti non progrediscono più tanto in quantità di subordinate impiegate, quanto in distribuzione di clausole per tipo e in

lunghezza. In modo analogo procede l'accuratezza: le apprendenti mostrano traiettorie di sviluppo a U, con riduzioni legate a sperimentazioni o ristrutturazioni di forme linguistiche più avanzate. Questo dato è in linea con quanto evidenziato in diversi studi acquisizionali (e.g. Larsen-Freeman, 1983) che mostrano come dal livello intermedio all'avanzato possa aumentare l'incidenza di errori, con la minor accuratezza accompagnata da evoluzioni interlinguistiche. D'altronde anche i modelli psicolinguistici (Skehan, 1998; cfr. § 2.1.2) mostrano come l'accuratezza sia determinata dall'interazione di tre elementi: il grado di accuratezza della rappresentazione stessa, la forza di eventuali rappresentazioni in competizione tra loro e infine il grado di automatizzazione della produzione stessa. In altre parole, in questo caso l'andamento non lineare è legato all'interazione tra i due costrutti, accuratezza e sviluppo, con apprendenti che possono essere molto accurati ma poco evoluti in una prima fase, e meno accurati ma più evoluti in una successiva, fino ad ottenere un riallineamento sui valori nativi. In questo caso con evoluto si intende sia la vicinanza della rappresentazione mentale alla norma della lingua target sia il grado di automatizzazione (cfr. Wolfe-Quintero *et al.*, 1998).

L'analisi della variazione longitudinale conferma dunque quanto già suggerito in letteratura, il progresso non è caratterizzato da un aumento in tutte le dimensioni contemporaneamente. Anzi, a livello avanzato il progresso può riguardare alcune dimensioni più di altre ed essere ottenuto sia tramite aumenti che riduzioni dei valori delle misure CAF.

### **Come variano le tre dimensioni CAF con il variare dell'attività comunicativa per gli apprendenti rispetto ai parlanti nativi?**

L'analisi della variazione situazionale effettuata disaggregando i dati di gruppo per task, permette di raffinare ulteriormente il quadro sopra delineato.

Alla prima rilevazione, le studentesse italiane dimostrano variazione tra task, nel senso che i loro task monologici sono sintatticamente più complessi, e un po' meno accurati e fluenti di quelli interattivi. Questo dato è in linea con i

modelli psicolinguistici dell'*information processing theory* secondo i quali i parlanti gestiscono le proprie risorse per rispondere alle richieste della situazione comunicativa, favorendo in alcuni casi la fluenza, in altri la complessità o l'accuratezza (cfr. Skehan, 1998; Robinson, 2005; 2006). Anche per le apprendenti c'è variazione tra task fin dalla prima rilevazione, con differenze poco evidenti rispetto alle italiane, a volte di polarità inversa. Nel primo caso si tratta dei due task monologici, nel secondo dei task interattivi. Qui le apprendenti risultano più complesse e fluenti delle italiane.

In prospettiva longitudinale tutte le ragazze registrano notevoli progressi, senza sorprese nel task in cui differivano maggiormente dalle italiane alla prima rilevazione. Gli andamenti nel tempo delle tre dimensioni variano con i diversi task, permettendo di identificare così una seconda ragione degli andamenti più o meno linearmente ascendenti delle misure considerate. Poiché i task monologici richiedono maggiore controllo sulla complessità, e quelli interattivi sulla fluenza, le apprendenti devono acquisire questo quadro di variazione. Pertanto nei task interattivi il progresso si realizza attraverso una riduzione della complessità; quando saper gestire un evento altamente interattivo come l'apertura delle telefonate di servizio, non richiede un mero controllo delle regole sintattiche, ma piuttosto la capacità di produrre una rapida serie di turni. Il tipo di task può quindi generare andamenti non lineari in virtù di un riallineamento con i pattern di variazione situazionale dei nativi. I miei dati dunque, in accordo con quanto evidenziato dalla letteratura con approccio sociolinguistico, evidenziano come apprendere una L2 significa acquisire anche i pattern di variazione tipici dei nativi (cfr. Rehner & Mougeon, 1999). Di qui l'importanza di un'osservazione della performance in diversi task e contesti comunicativi, con la variabile monologico-interattivo che si rivela particolarmente fruttuosa.

### **Come si sviluppa il percorso individuale di ciascun soggetto?**

L'analisi dei percorsi individuali mostra variazione individuale per tutti i soggetti. Le studentesse italiane registrano variazione individuale soprattutto nei task monologici. Alla prima rilevazione Elisa produce narrazioni più snelle e meno complesse della compagna, risultando pertanto più fluente e accurata. Questo dato conferma quanto suggerito dai modelli psicolinguistici. Un parlante infatti dovendo gestire le proprie risorse d'attenzione che sono limitate, nel realizzare un racconto può decidere di volta in volta se adottare uno stile narrativo che favorisce la complessità, a scapito di accuratezza e fluenza o viceversa. Nel tempo le due apprendenti convergono su valori simili. I cambiamenti registrati da Valentina nel task film permettono un'ulteriore riflessione: la variazione individuale può essere legata a fattori esterni sia al task che alla competenza. In questo caso infatti Valentina produce una narrazione semplice, frammentaria e poco fluente perché incontra alcune difficoltà nella comprensione del video e si mostra poco coinvolta nell'attività.

Anche le apprendenti mostrano variazione individuale, Pandita e Catherine al t1 risultano meno avanzate di Eden e Shirley. Per le prime due ragazze il progresso nel corso dei tre anni è prevedibilmente maggiore che per le seconde due. L'analisi mostra poi come l'avvicinamento al comportamento linguistico delle studentesse italiane procede in modo e in misura diversa da apprendente a apprendente. Questo è dovuto in alcuni casi a scelte individuali, è il caso di Shirley, una delle apprendenti più avanzate. Tra t1 e t4 Shirley nei task monologici riduce la complessità, cambiando stile narrativo, scegliendo una narrazione caratterizzata da clausole indipendenti e frammenti giustapposti sostenuti da un'intonazione calzante ed espressiva. In altri casi ciò dipende dal livello alla prima rilevazione. Le apprendenti meno avanzate nei task monologici aumentano la complessità e seguono una traiettoria di sviluppo a U in accuratezza, nei task interattivi invece si riduce la complessità, a favore di accuratezza e fluenza. Le apprendenti più avanzate riducono la complessità in tutti i task, con progressi tendenzialmente più lineari in accuratezza e fluenza.



Dati di gruppo e dati individuali risultano così complementari: il gruppo definisce le tendenze generali, il caso individuale osservato longitudinalmente permette di ricostruire ciò che accade durante il percorso di apprendimento. Importante sottolineare come nei miei dati la variazione individuale si collochi in genere entro lo stesso quadro di variazione del gruppo, sostenendo così l'idea suggerita dalla letteratura (cfr. Pienemann, 2008; Larsen-Freeman, 2006), secondo cui l'interlingua procede per traiettorie individuali e aderisce contemporaneamente a una sequenzialità più universale.

### **Quali sono le interazioni tra queste tre dimensioni dell'interlingua?**

Le tre dimensioni CAF possono avere traiettorie di sviluppo non lineari, oltre che per i fattori visti sopra, in virtù delle interazioni che si creano tra loro. È questo il caso delle produzioni delle studentesse italiane e di Pandita e Catherine nei task monologici. Per le studentesse italiane, le scelte stilistiche individuali che abbiamo appena commentato rispetto alle modalità narrative generano un trade-off tra le tre dimensioni. Se realizzare un racconto complesso significa ridurre l'attenzione alla forma e dunque all'accuratezza, realizzarne uno più semplice e fluente, porta a una riduzione della complessità. In accordo con i modelli psicolinguistici dell'*information processing view* (Skehan, 1998; cfr. § 2.2.1.2), nello svolgimento di un compito comunicativo il parlante può gestire le proprie risorse, e scegliere in alcuni casi di favorire il significato affidandosi al sistema basato sugli esempi, con un conseguente aumento della fluenza, in altri di favorire il sistema basato sulle regole, ottenendo così una maggiore complessità. Per Pandita e Catherine, la produzione di narrazioni più complesse e l'uso di forme di subordinazione più avanzate generano in alcune fasi compensazione sull'accuratezza e sulla fluenza. Questo effetto non si realizza invece nei task interattivi, proprio perché questi ultimi richiedono la produzione di AS-unit semplici e rapide, favorendo pertanto uno sviluppo più lineare di accuratezza e fluenza.

L'andamento non lineare di alcune misure, dovuto a una serie di fattori di natura alquanto diversa – la diversa competenza linguistica delle apprendenti alla prima rilevazione, il tipo di task che svolgono, lo stile discorsivo che adottano, e infine le interazioni tra dimensioni, cui conseguono effetti di trade-off – suggerisce dunque una nota di cautela nel generalizzare conclusioni sulla competenza in base a un unico dato, poiché a un aumento o a una diminuzione in una misura non corrisponde automaticamente un progresso o un regresso. Di qui l'importanza di una loro combinazione e di un confronto con la variazione nativa quando l'obiettivo è arrivare all'elaborazione di profili di proficiency più affidabili.

**I descrittori selezionati dal QCER sono efficaci nella valutazione delle competenze orali di studentesse italiane e apprendenti di italiano L2 di livello avanzato? In particolare sono sensibili allo sviluppo longitudinale e alla variazione tra task?**

Nel complesso i descrittori del QCER si sono rivelati efficaci nella valutazione delle competenze orali tanto delle studentesse italiane quanto delle apprendenti e hanno rilevato variazione longitudinale e situazionale.

Le italiane, pur non ottenendo sempre il punteggio massimo, si collocano nella fascia C, e vengono dunque valutate come più avanzate delle apprendenti. Queste ottengono valutazioni più basse in tutti i task, con differenze più evidenti per Pandita e Catherine, che sono peraltro meno avanzate di Eden e Shirley. In prospettiva longitudinale, le apprendenti dimostrano progressi, registrando nel corso dei tre anni, seppur in misura diversa e relativamente a descrittori diversi da studentessa a studentessa, il passaggio dal livello B al livello C.

Nonostante l'efficacia complessiva dei descrittori QCER, sono emerse alcune difficoltà, in particolare rispetto al descrittore coerenza-coesione per i task interattivi. Questo evidenzia come il QCER non sia sempre sensibile alla

variazione tra task. Non a caso infatti, a questo proposito le discussioni tra valutatori per raggiungere un accordo sono state più lunghe e a volte infruttuose. Come già evidenziato nelle recenti pubblicazioni in merito (per es. Alderson *et al.*, 2004; Weir, 2005), il QCER nella sua formulazione attuale non è sempre in grado di rispondere ai bisogni concreti di chi si occupa di valutazione. Chi utilizza le griglie avverte la necessità di sapere qual è il livello di qualità della performance che ci si aspetta nello svolgere un determinato compito a un determinato livello, come nel caso delle telefonate. Superare questi limiti richiederà di approfondire il legame tra i diversi livelli e le forme d'uso, in particolare per i livelli più alti, che pongono appunto le maggiori difficoltà (cfr. Alderson *et al.*, 2004; Ferrari & Nuzzo, in stampa). Su questo tema è auspicabile una maggiore interazione tra la linguistica acquisizionale e il testing.

### **Che relazione intercorre tra le misure analitiche CAF e le valutazioni soggettive della proficiency basate sul QCER?**

Il confronto tra i risultati ottenuti con le misure CAF e quelli ottenuti con le valutazioni QCER ha permesso di verificare come i due sistemi di misurazione rilevano nel complesso tendenze di variazione longitudinale e situazionale tra loro coerenti. In alcuni casi essi disegnano le stesse tendenze, in altri casi invece l'interpretazione di uno complementa quella dell'altro.

Si ritrovano le stesse tendenze nella variazione individuale nel tempo, in entrambe le misurazioni infatti le studentesse italiane e due delle apprendenti risultano più avanzate, con conseguenti differenze nei percorsi di sviluppo. Si tratta invece di interpretazioni complementari tra CAF e QCER nel caso degli andamenti non lineari delle misure CAF per complessità e accuratezza. Da un lato, il QCER può guidare l'interpretazione di CAF. Infatti abbiamo visto come il tipo di task, monologico o interattivo, il livello di partenza, le scelte individuali del parlante e il trade-off tra dimensioni CAF generano andamenti

non sempre lineari nelle misure considerate, mentre il QCER registra ovunque progresso, guidando pertanto l'interpretazione degli andamenti delle misure stesse. Dall'altro lato, è invece il CAF che potrebbe contribuire a raffinare la valutazione del QCER. L'analisi CAF mostra come i task interattivi favoriscano le produzioni di AS-unit semplici (ossia con subordinazione ridotta e limitata ad alcune forme specifiche), con una bassa densità informativa e l'introduzione di un elemento per volta. Questi elementi potrebbero costituire un'utile base per una ridefinizione del descrittore coerenza-coesione, che risulta invece poco adatto a rilevare questa variazione situazionale.

## *5.2. Implicazioni per la ricerca acquisizionale, il testing e la didattica*

I risultati ottenuti in questo lavoro, per quanto riferiti ad un campione limitato e pertanto da considerarsi con cautela, ci permettono di trarre alcune implicazioni per la ricerca acquisizionale, per il testing e per la didattica, in parte confermando quanto già evidenziato da altri, in parte contribuendo qualcosa di nuovo.

Per quanto riguarda l'acquisizione, i nostri risultati confermano ciò che è emerso anche in studi precedenti, ovvero che alcuni aspetti dell'interlingua variano sistematicamente secondo i task e le situazioni comunicative, e che pertanto ogni task fornisce un'immagine parziale dell'interlingua. Molti studi acquisizionali invece traggono le loro conclusioni usando dati elicitati per mezzo di un solo tipo di compito, spesso monologico. Una ricerca come la nostra mostra che attraverso la combinazione di più task è possibile creare profili più attendibili delle competenze di un parlante e avere informazioni più precise circa gli andamenti nel tempo. La nostra analisi ha evidenziato inoltre come il percorso evolutivo non corrisponda sempre ad un aumento lineare e generalizzato della complessità sintattica, dell'accuratezza o della fluenza.

Per quanto riguarda il testing linguistico, i risultati di questo studio sottolineano che per ottenere una valutazione precisa dello sviluppo della

proficiency è necessario osservare le produzioni degli apprendenti in diverse attività e contesti, rappresentativi di tutte le competenze che si vogliono testare. I dati evidenziano inoltre che è utile non separare rigidamente la competenza linguistica dalla competenza interazionale, proprio perché esse sono strettamente connesse. Infine lo studio conferma l'efficacia di una valutazione basata sui task, come sostengono ormai molti autori (Brown *et al.*, 2002; *Language Testing*, 2002). L'uso dei task, oltre a garantire una maggiore validità esterna o ecologica, consente appunto di esaminare come la competenza linguistica e quella comunicativa interagiscono nello svolgimento di specifiche attività.

Per quanto riguarda l'interfaccia tra misure CAF e valutazioni QCER, l'indagine svolta conferma l'utilità in questo senso di una maggiore interazione tra ricerca acquisizionale e testing linguistico. Chi si occupa di testing può trovare nell'acquisizione spunti per una migliore definizione degli elementi linguistici corrispondenti ai livelli di proficiency, necessari per un maggior affinamento delle scale di proficiency. D'altro canto, chi si occupa di acquisizione può trovare nel testing strumenti per una più rigorosa verifica dell'affidabilità degli strumenti di elicitazione e misurazione, elemento indispensabile per ottenere non solo risultati più robusti, ma permettere anche una maggior replicabilità.

Per l'insegnamento, nella costruzione di un syllabo linguistico-comunicativo la variabilità situazionale riscontrata nella nostra analisi mostra quanto è importante tener conto dei dettagli dell'interazione e delle routine ad esse collegate. Ma per conoscerli, come nel caso dell'apertura della telefonata, non basta ricorrere all'intuizione dell'insegnante, ma è doveroso partire dai dati della ricerca (Bettoni, 2006: 237). Anche sul piano della didattica occorre dunque approfondire il legame tra competenza linguistica e competenza interazionale. I miei dati mostrano che esiste una sintassi sui generis per l'interazione, tale che la glottodidattica dovrà tenere conto delle ricerche sulla *grammar in interaction* (Ochs *et al.*, 1996; Ford *et al.*, 2002), contestualizzando l'uso delle varie forme linguistiche all'interno di pratiche

comunicative. Fermo restando che l'obiettivo generale dovrebbe essere quello di stimolare negli apprendenti l'attenzione e la consapevolezza per certi fenomeni, l'apertura di una telefonata può costituire un buon esempio di situazione didattizzabile: poche mosse, relativamente predeterminate quanto al contenuto, possono essere realizzate in una grande varietà di modi, ciascuno dei quali avrà conseguenze sui dettagli della dinamica interazionale di un particolare frammento (cfr. Pallotti & Ferrari, 2008). Si aprono così interessanti prospettive per una didattica della grammatica focalizzata sui micro-dettagli dell'interazione, che può condurre all'elaborazione di sillabi linguistico-comunicativi innovativi.



## Riferimenti bibliografici

- ADAMSON H.D. & REGAN V. (1991), The acquisition of community norms by Asian immigrants learning English as a second language: A preliminary study, in *Studies in Second Language Acquisition* 13, 1-22.
- ALDERSON J.C. (1991), Bands and scores, in J.C. Alderson & B. North (eds.), *Language Testing in the 1990s*, Macmillan, London, 71-86.
- ALDERSON J.C. (2002) (ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case studies*, Council of Europe, Strasbourg.
- ALDERSON J.C. (2005), Editorial, in *Language Testing* 22, 257-60.
- ALDERSON J.C., CLAPHAM C. & WALL D. (1995), *Language test construction and evaluation*, Cambridge University Press, Cambridge.
- ALDERSON J.C., FIGUERAS N., KUIJPER H., NOLD G., TAKALA S. & TARDIEU C. (2004), *The development of specifications for item development and classification within the Common European Framework of Reference for Languages, learning, teaching, assessment. Reading and listening: final report of the Dutch Construct Project*. Available on request from the Project Coordinator, J. Charles Alderson, c.alderson@lancaster.ac.uk.
- ALDERSON J.C. & URQUHART A.H. (1985), The effect of students' academic discipline on their performance on ESP reading tests, in *Language Testing* 2, 192-204.
- ANDERSON J.R. (1995), *Learning and Memory: An Integrated Approach*, Wiley, New York.
- ANDERSON N., BACHMAN L.F., COHEN A.D. & PERKINS K. (1991), An exploratory study into the construct validity of a reading comprehension test, triangulation of data sources, in



- Language Testing* 8, 41-66.
- APPLIED LINGUISTICS (2006), *Special Issue* 27, G. Cook & G. Kasper (eds.).
- ATKINSON J.M. & HERITAGE J. (1984), Transcription Notation, in J.M. Atkinson & J. Heritage (eds.), *Structure of Social Interaction: Studies in Conversation Analysis*, Cambridge University Press, Cambridge, ix-xvi.
- AURELI M. (2002), Pressione dell'uso sulla norma. Le relative non standard nei giudizi degli utenti, Preprint 13, 1-20, <http://lablita.dit.unifi.it/italian/preprint.html>
- BACHMAN L.F. (1988), Language testing-SLA research interfaces, in *Annual Review of Applied Linguistics* 9, 193-209, ristampato in L.F. Bachman & A. Cohen (1998) (eds.).
- BACHMAN L.F. (1990), *Fundamental considerations in language testing*, Oxford University Press, Oxford.
- BACHMAN L.F. (1998), Language testing-SLA research interfaces, in L.F. Bachman & A.D. Cohen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, 177-95.
- BACHMAN L.F. (2000), Modern Language Testing at the turn of the century: assuring that what we count counts, in *Language Testing* 17, 1-41.
- BACHMAN L.F. (2004), *Statistical analysis for language assessment*, Cambridge University Press, Cambridge.
- BACHMAN L.F. & COHEN A. (1998a) (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge.
- BACHMAN L.F. & COHEN A. (1998b), Language testing-SLA interfaces: An update, in L.F. Bachman & A. Cohen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, 1-31.
- BACHMAN L.F., DAVIDSON F. & MILANOVIC M. (1996), The use of test methods in the content analysis and design of EFL proficiency tests, in *Language Testing* 13, 125-50.
- BACHMAN L.F. & PALMER A.S. (1981), A multitrait multimethod investigation into the construct validity of six tests of speaking and reading, in A.S. Palmer, J.M. Groot & G.A. Trosper (eds.), *The construct validation of tests of communicative competence*, TESOL, Washington, 149-63.
- BACHMAN L.F. & PALMER A.S. (1996), *Language testing in practice*, Oxford University Press, Oxford.
- BACHMAN L.F. & SAVIGNON S.J. (1986), The evaluation of communicative language proficiency: A critique of the ACTFL oral interview, in *The Modern Language Journal* 70, 380-90.
- BADDELEY A. (2007), *Working Memory, Thought and Action*, Oxford University Press, Oxford.
- BANERJEE J., FRANCESCHINA F. & SMITH A. M. (2007), Documenting features of written

- language production typical at different IELTS band score levels, in *IELTS Research Report 7*.
- BARDOVI-HARLIG K. (1992), A second look at T-unit analysis: Reconsidering the sentence, in *TESOL Quarterly* 26, 390-5.
- BARDOVI-HARLIG K. & BOFMAN T. (1989), Attainment of syntactic and morphological accuracy by advanced language learners, in *Studies in Second Language Acquisition* 11, 17-34.
- BAYLEY C.J.N. (1973), *Variation and Linguistic Theory*, Center for Applied Linguistic, Washington D.C.
- BAYLEY R. (1991), *Variation theory and second language learning: Linguistic and social constraints on interlanguage tense marking*, PhD. Dissertation, Stanford University, citato in Preston & Bayley (1996).
- BAYLEY R. (1994), Interlanguage variation and the quantitative paradigm: Past-tense marking in Chinese-English, in E. Tarone, S. Gass & A. Cohen (eds.), *Research methodology in Second Language Acquisition*, Lawrence Erlbaum, Hillsdale, NJ, 157-81.
- BAYLEY R. (1996), Competing constraints on variation in the speech of adult Chinese learners of English (eds.), in R. Bayley & D. Preston, *Second Language Acquisition and Linguistic Variation*, John Benjamins, Amsterdam, 96-120.
- BAYLEY R. & LANGMAN J. (2004), Variation in the group and the individual: Evidence from second language acquisition, in *IRAL International Review of Applied Linguistics* 42, 303-18.
- BAYLEY R. & PRESTON D. (1996) (eds.), *Second Language Acquisition and Linguistic Variation*, John Benjamins, Amsterdam.
- BAZZANELLA C. (1991), Tratti substandard nel parlato colloquiale, in C. Lavinio & A. Sobrero, *La lingua degli studenti universitari*, La Nuova Italia, Firenze, 123-63.
- BAZZANELLA C. (1994), *Le facce del parlare. Un approccio pragmatico all'italiano parlato*, La Nuova Italia, Firenze.
- BAZZANELLA C. (2002) (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerrini e Associati, Milano.
- BEATTIE G. (1980), The role of language production in the organisation of behaviour in face-to-face interaction, in B. Butterworth (ed.), *Language production: Vol I*, Academic Press, London, 69-109.
- BEEBE L.M. (1981), Social and Situational Factors Affecting the Communicative Strategy of Dialect Code-Switching, in *International Journal of the Sociology of Language* 31, 139-49.
- BEEBE L.M. (1985), *Reservations about the Labovian paradigm of style shifting and its extension to the study of interlanguage*, ERIC document.

- BEEBE L.M. (1987), Sociolinguistic variation and style shifting in second language acquisition, in *Language Learning*, 2, 1980, 433-48, reprinted in G. Ioup & S. Weinberger (eds.), *Interlanguage Phonology*, Newbury House Publishers.
- BEEBE L.M. & ZUENGLER J. (1983), Accommodation theory: An explanation for style shifting in second language dialects, in N. Wolfson & E. Judd (eds.), *Sociolinguistics and language acquisition*, Newbury, Rowley MA, 195-213.
- BERCELLI F. & PALLOTTI G. (2002) Conversazioni telefoniche, in C. Bazzanella (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerini, Milano, 177-192.
- BERNINI G. (1991), Frasi relative nel parlato colloquiale, in C. Lavinio & A. Sobrero, *La lingua degli studenti universitari*, La Nuova Italia, Firenze, 165-87.
- BERRUTO G. (1986), Un tratto sintattico dell'italiano parlato: il c'è presentativo, in K. Lichem, E. Mara & S. Knaller (a cura di), *Parallela 2. Aspetti della sintassi dell'italiano contemporaneo*, Tübingen, Narr, 61-73.
- BERRY V. (1993), Personality characteristics as a potential source of language test bias, in A. Huhta, K. Sajavaara & S. Takala (eds.), *Language testing: new openings*, University of Jyväskylä, Jyväskylä, 114-24.
- BERWICK R. & ROSS S. (1996), Cross-cultural pragmatics in oral proficiency interview strategies, in M. Milanovic & N. Saville (eds.), *Performance testing, cognition and assessment*, University of Cambridge Local Examinations Syndicate and Cambridge University Press, Cambridge, 34-54.
- BETTONI C. (2006), *Usare un'altra lingua. Guida alla pragmatica interculturale*, Laterza, Bari-Roma.
- BIBER D., CONRAD S. & REPPEN R. (1998), *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, Cambridge.
- BICKERTON D. (1971), Inherent variability and variable rules, in *Foundations of Language* 7, 457-92.
- BILINGUALISM: LANGUAGE AND COGNITION (2007), Special Issue 10, 1.
- BOSISIO C. (2005), *Dagli approcci tradizionali al Quadro comune europeo di riferimento. Riflessioni glottodidattiche e applicazioni per l'insegnante di italiano L2*, ISU Università Cattolica, Milano.
- BREEN M. (1987), Contemporary paradigms in syllabus design (Parts 1 and 2), in *Language Teaching* 20, 91-2 and 157-74.
- BRINDLEY G. (1998), Describing language development? Rating scales and SLA. In L.F. Bachman & A.D. Cohen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, 112-40.
- BROWN A. (1995), The effect of rater variables in the development of an occupation-specific language performance test, in *Language Testing* 12, 1-15.

- BROWN A. (2003), Interviewer variation and the co-construction of speaking proficiency, in *Language Testing* 20, 1-25.
- BROWN A. (2005), *Interviewer variability in language proficiency interviews*, Peter Lang, Frankfurt.
- BROWN J.D., HUDSON T.D., NORRIS J.M. & BONK W. (2002), *An investigation of second language task-based performance assessments*, University of Hawaii Press, Honolulu, HI.
- BROWN R. (1973), *A First Language: the Early Stages*, Harvard University Press, Cambridge, Mass.
- BRUMFIT C. (1984), *Communicative Methodology in Language Teaching*, Cambridge University Press, Cambridge.
- BRUMFIT C. (1988), Applied linguistics and communicative language teaching, in W. Grabe (ed.) *Annual Review of Applied Linguistics: Communicative Language Teaching*, Cambridge University Press, New York, 1988.
- BRUMFIT C. & JOHNSON K. (1979), *The Communicative Approach to Language Teaching*, Cambridge University Press, New York.
- BRUNER J.S. (1973), *The relevance of education*, Norton, New York.
- BUCK G. (1991), The testing of listening comprehension: an introspective study, in *Language Testing* 8, 67-91.
- BYGATE M. (1996), Effects of task repetition: appraising the developing language of learners, in J. Willis & D. Willis (eds.), *Challenge and Change in Language Teaching*, Heinemann, London, 136-46.
- BYGATE M. (1999), Task as context for framing, reframing and unframing language, in *System* 27, 33-48.
- BYGATE M. (2001), Effects of task repetition on the structure and control of oral language, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic tasks, Second Language Learning, Teaching and Testing*, Longman, London, 23-48.
- BYGATE M., DÖRNYEI Z., CSÖLLE A., KIRÁLY K., KORMOS J. & NEMETH N. (1998), *Perspective from a task-based instruction project*, Colloquium, 32nd Annual TESOL Convention, Seattle.
- BYGATE M., SKEHAN P. & SWAIN M. (2001) (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, Longman, London, 2001.
- CAMERON L. & DEIGNAN A. (2006), The emergence of metaphor in discourse, in *Applied Linguistics*, Special Issue 27, 671-90.
- CANALE M. (1983), On some dimensions of language proficiency, in J.W. Oller (ed.), *Issues in language testing research*, Newbury House, Rowley, Mass., 333-42.
- CANALE M. & SWAIN M. (1980), Theoretical basis of communicative approaches to second

- language teaching and testing, in *Applied Linguistics* 1, 1-47.
- CANALE M. & SWAIN M. (1981), A theoretical framework for communicative competence, in A.S. Palmer, J.M. Groot & G.A. Trosper (eds.), *The construct validation of tests of communicative competence*, TESOL, Washington, 31-6.
- CANDLIN C.N. (1987), Towards task-based language learning, in C.N. Candlin & D.F. Murphy, *Language Learning Tasks*, in Lancaster Practical Papers in English Language Education vol. 7, Prentice Hall, Cambridge, 5-22.
- CARROLL J.B. (1993), *Human Cognitive Abilities*, Cambridge, New York.
- CARROL J.B. & WEST R. (1989), *ESU Framework. Performance Scales for English Language Examinations*, Longman, Harlow.
- CASSANDRO M. & MAGGINI M. (2004). Osservazione e valutazione di apprendenti stranieri in relazione ai livelli comuni di riferimento del Quadro comune europeo, in E. Jafrancesco (a cura di), *Le tendenze innovative del quadro comune europeo di riferimento per le lingue ed il portfolio. Atti del XII convegno nazionale ILSA*. Edilingua, Firenze, 64-107.
- CHALHOUB-DEVILLE M. (1995a), Deriving oral assessment scales across different tests and rater groups, in *Language Testing* 12, 16-33.
- CHALHOUB-DEVILLE M. (1995b), A contextualized approach to describing oral language proficiency, in *Language Learning* 45, 251-81.
- CHALHOUB-DEVILLE M. (1996), Performance assessment and the components of the oral construct across different tests and rater groups, in M. Milanovic & N. Saville (eds.), *Performance testing, cognition and assessment*, University of Cambridge Local Examinations Syndacate and Cambridge University Press, Cambridge, 55-73.
- CHALHOUB-DEVILLE M. (2001), Task-based assessments: Characteristics and validity evidence, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, 210-28.
- CHINI M. (1998), La subordinazione in testi narrativi di apprendenti tedescofonici: forma e funzione, in *Linguistica e filologia* 7, 121-59.
- CHINI M. (2003), Le phénomène de la jonction interpropositionnelle dans la narration en italien L2: agrégation et intégration, in *AILE Acquisition et Interaction en Langue Etrangère* 17, 71-106.
- CHOMSKY N. (1965), *Aspects of the Theory of Syntax*, The MIT Press, Cambridge (Mass.).
- CHOMSKY N. (1975), *Reflections on Language*, Temple Smith, London.
- CHOMSKY N. (1980), *Rules and Representations*, Blackwell, Oxford.
- CINQUE G. (2001), La frase relativa, in L. Renzi & G. Salvi, *Grande grammatica italiana di consultazione*, vol. I. *La frase. I sintagmi nominale e preposizionale*, Il Mulino, Bologna, 443-503.
- CLAHSEN H. (1980), Psycholinguistic aspects of L2 acquisition, in S.W. Felix (ed.), *Second*

- language development – trend and issue*, Tübingen, 57-79.
- CLAHSEN H. & MUYSKEN P. (1986), The availability of UG to adult and child learners. A study of the acquisition of German word order, in *Second Language Research* 2, 93-119.
- CLAHSEN H. & MUYSKEN P. (1989), The UG paradox in L2 acquisition, in *Second Language Research* 5, 1-29.
- CLAPHAM C. (1993), The effect of language proficiency and background knowledge on EAP students' reading comprehension, in A.J. Kunnan (ed.), *Validation in language assessment*, Lawrence Erlbaum, Mahwah, 141-68.
- CLAPHAM C. (1996), *The development of IELTS: a study of the effect of background knowledge on reading comprehension*, University of Cambridge Local Examinations Syndicate and Cambridge University Press, Cambridge.
- CONSIGLIO D'EUROPA (1988), *Final report of the project report of project 12: learning and teaching modern languages for communication*, Council for Cultural Co-operation, Strasbourg, CC-GP12, 11.
- CONSIGLIO D'EUROPA (1992), *Transparency and coherence in language learning in Europe. Objectives, evaluation, certification*. Report on the Rüschlikon Symposium.
- CONSIGLIO D'EUROPA (1996), *Modern languages: learning, teaching, assessment: a Common European Framework of Reference*. Draft 2 of a framework proposal, Council for Cultural Co-operation, Education Committee, Strasbourg.
- CONSIGLIO D'EUROPA (2001), *Quadro comune europeo di riferimento per le lingue*, La Nuova Italia, Scandicci, Firenze.
- CONSIGLIO D'EUROPA (2002), *Relating language examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment* (CEF), DGIV/EDU/LANG 2002, 15, Language Policy Division, Strasbourg.
- CONSIGLIO D'EUROPA (2003), *Relating language examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment* (CEF). Manual: Preliminary Pilot Version. DGIV/EDU/LANG 2003, 5, Language Policy Division, Strasbourg.
- CONSIGLIO D'EUROPA (2005), *Reference supplement to the preliminary version of the manual for relating examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment*, DGIV/EDU/LANG 2005, 13, Language Policy Division, Strasbourg.
- COOPER T.C. (1976), Measuring written syntactic patterns of second language learners of German, in *The Journal of Educational Research* 69, 176-83.
- CORDER S. P. (1967), The significance of learners' errors, in *IRAL International Review of Applied Linguistics* 5, 161-9.
- CORDER S. P. (1981), *Error analysis and interlanguage*, Oxford University Press, Oxford.

- CRESTI E. (1987), *L'articolazione dell'informazione nel parlato*, in A.A.V.V., *Gli italiani parlati*, Accademia della Crusca, Firenze, 27-90.
- CRESTI E. (1992a), Definizione dell'enunciato e pragmatica, in L. Brasca & M.L. Zambelli (a cura di), *Grammatica dell'ascoltare e del parlare a scuola*, Atti del V° Convegno nazionale GISCEL, La Nuova Italia, Firenze.
- CRESTI E. (1992b), La scansione del parlato e l'interpunzione, in E. Cresti, N. Maraschio & L. Toschi (a cura di), *Storia e teoria dell'interpunzione*, Bulzoni, Roma, 443-99.
- CRESTI E. (1993), Formazione dell'enunciato e articolazione topic-comment nominale, in *Studi di grammatica italiana*, Accademia della Crusca, vol. XIV, Firenze.
- CRESTI E. (1994), *Information and intonational patterning in Italian*, in B. Ferguson, H. Gezundhajt & P. Martin (eds.) *Accent, intonation et modèles phonologiques*, Editions Mélodie, Toronto.
- CRESTI E. (1995), *Speech acts units and informational units*, in E. Fava (a cura di), *Speech acts and linguistic research, Proceedings of the Workshop*, July 15-17, 1994, Center for Cognitive Science, State University of New-York at Buffalo, Edizioni Nemo, Padova, 89-107.
- CRESTI E. (1998), *Gli enunciati nominali*, in M.T. Navarro (a cura di), Atti del IV° Convegno Internazionale SILFI (Madrid), Cesati, Pisa, 171-91
- CRESTI E. (2001), Per una nuova definizione di frase, in P. Bongralli, A. Dardi, M. Fanfani & R. Tesi (a cura di), *Studi di storia della lingua italiana offerti a Ghino Ghinassi*, Le Lettere, Firenze, 511-50.
- CRESTI E. (2005), *Enunciato e frase: teoria e verifiche empiriche*, in M. Biffi, O. Calabrese & L. Salibra (a cura di), *Italia Linguistica: discorsi di scritto e di parlato. Scritti in onore di Giovanni Nencioni*, Prolagon, Siena, 249-60.
- CRESTI E. (2005), Per una nuova classificazione dell'illocuzione, in E. Burr (a cura di), *Tradizione e innovazione*, Atti del VI convergno della Società di Linguistica e Filologia Italiana (SILFI), Cesati, Firenze, 233-46.
- CRESTI E. & GRAMIGNI P. (2004), Per una linguistica corpus based dell'italiano parlato: le unità di riferimento, in F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy (a cura di), *Il parlato italiano*, Atti del Convegno Nazionale, D'Auria Editore, Napoli, 1-23.
- CRESTI E. & MONEGLIA M. (2005) (a cura di), *C-ORAL-ROM, Integrated Reference Corpora for Spoken Romance Language*, full edition in 9DVDs, ELRA, Paris.
- CRONBACH L.J. (1971), Test validation, in R.L. Thorndike (ed.), *Educational Measurement*. 2nd edition, American Council on Education, 443-507.
- CRONBACH L.J. & MEEHL P.E. (1955), Construct validity in psychological tests, in *Psychological Bulletin* 52, 281-302.
- CROOKES G. (1986), *Task classification: A cross-disciplinary review*, Technical Report 4,

- Center for Second Language Classroom Research, University of Hawaii, Honolulu.
- CROOKES G. (1989), Planning and interlanguage variation, in *Studies in Second Language Acquisition* 11, 367-83.
- CROOKES G. (1990), The utterance and other basic units for second language discourse analysis, in *Applied Linguistics* 11, 183-199.
- CROOKES G. & GASS S. (1993) (eds.) *Tasks in language learning: Integrating theory and practice*, Multilingual Matters, Clevedon, Avon.
- DAVIES A. (1990), *Principles of language testing*, Basil Blackwell, London.
- DAVIES A. (2003), Three heresis of language testing research, in *Language Testing* 20, 355-68.
- DE BOT K., LOWIE W. & VERSPOOR M. (2007), A dynamic systems theory to second language acquisition, in *Bilingualism: Language and Cognition* 10, 7-21
- DE CAMP D. (1973), Implicational scales and sociolinguistic linearity, in *Linguistics* 73, 30-43.
- DE JONG J. (1992), Assessment of language proficiency in the perspective of the 21st century, in *AILA Review* 9, 39-45.
- DE MAURO T. & THORTON A. M. (1985), *La predicazione: teoria e applicazione all'italiano*, in A. Franchi-De Bellis e L.M. Savoia (a cura di), *Sintassi e morfologia della lingua italiana d'uso*, Bulzoni, Roma.
- DEWAELE J.M. (1992), L'omission du *ne* dans deux styles oraux d'interlangue française, in *Interface: Journal of Applied Linguistics* 7, 3-17.
- DEWAELE J.M. & REGAN V. (2002), Maîtriser la norme sociolinguistique en interlangue français: le cas de l'omission variable de *ne*, in *Journal of French Language Studies* 12, 123-48.
- DICKERSON L.J. (1975), The learners' interlanguage as a system of variable rules, in *TESOL Quarterly* 9, 401-7.
- DOUGLAS D. (2001), Performance consistency in second language acquisition and language testing research: A conceptual gap, in *Second Language Research* 17, 442-56.
- ELDER C. & IWASHITA N. (2005), Planning for test performance: does it make a difference?, in R. Ellis (ed.), *Planning and Task Performance in a Second Language*, Benjamins, Amsterdam, 219-37.
- ELLIS N.C. (1996), Sequencing in SLA: Phonological memory, chunking and points of order, in *Studies in Second Language Acquisition* 18, 91-126.
- ELLIS N.C. (1998), Emergentism, connectionism and language learning, in *Language Learning* 48, 631-64.
- ELLIS N.C. (2006a), Language acquisition as rational contingency learning, in *Applied Linguistics* 27, 1-24.
- ELLIS N.C. (2006b), Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning,



- in *Applied Linguistics* 27, 1–31.
- ELLIS N.C. (2006c), Cognitive perspectives on SLA: The Associative Cognitive CREED, in *AILA Review* 19, 100–21.
- ELLIS N.C. & LARSEN-FREEMAN D. (2006), Language emergence: Implications for applied linguistics (introduction to the special issue), in *Applied Linguistics* 27, 558–89.
- ELLIS R. (1985a), Sources of variability in interlanguage, in *Applied Linguistics* 6, 118–31.
- ELLIS R. (1985b), *Understanding second language acquisition*, Oxford University Press, Oxford.
- ELLIS R. (1987a), Interlanguage variability in narrative discourse: Style shifting in the use of past tense, in *Studies in Second Language Acquisition* 9, 1–20.
- ELLIS R. (1987b), Contextual variability in second language acquisition and the relevancy of language teaching, in R. Ellis (ed.), *Second Language Acquisition in Context*, Prentice-Hall International, London, 179–94.
- ELLIS R. (1988), The effects of linguistic environment on the second language acquisition of grammatical rules, *Applied Linguistics* 9, 257–74.
- ELLIS R. (1989), Sources of intra-learner variability on language use and their relationship to second language acquisition, in S. Gass, C. Madden, D. Preston, & L. Selinker (eds.), *Variation in Second Language Acquisition. Volume 2: Psycholinguistic Issues*, Multilingual Matters, Clevedon, 22–45.
- ELLIS R. (1990), Grammaticality judgements and learner variability, in H. Burmeister & P. Rounds (eds.), *Variability in Second Language Acquisition*, Oregon University, Oregon, 25–60.
- ELLIS R. (1994), *The study of second language acquisition*, Second edition, Cambridge University Press, Cambridge.
- ELLIS R. (1996), SLA and language pedagogy, in *Studies in Second Language Acquisition* 19, 69–92.
- ELLIS R. (1999), Item versus system learning: Explaining free variation, *Applied Linguistics* 20, 460–80.
- ELLIS R. (2000), Task-based research and language pedagogy, in *Language Teaching Research* 4, 193–220.
- ELLIS R. (2001), Some thoughts on testing grammar: An SLA perspective, in C. Elder, A. Brown, N. Iwashita, E. Grove, K. Hill, T. Lumley, T. McNamara & K. O'Loughlin (eds.), *Experimenting with uncertainty. Essays in honour of Alan Davies*, Cambridge University Press, Cambridge, 251–63.
- ELLIS R. (2003), *Task-based Language Learning and Teaching*, Oxford University Press, Oxford.

- ELLIS R. (eds.) (2005), *Planning and task performance in a second language*, Benjamins, Amsterdam.
- ELLIS R. & BARKHUIZEN G. (2005), *Analysing learner language*, Oxford University Press, Oxford.
- FERRARI S. & NUZZO E. (in stampa), La valutazione delle competenze orali in italiano L2. Una verifica sperimentale dei criteri suggeriti dal Quadro Comune Europeo, in *Atti del XV Convegno Nazionale Giscel*, Milano, 6-8 marzo 2008.
- FERRARIS S. (2001), Text organization in Italian L2 varieties, in S.H. Foster-Cohen & A. Nizgorodcew (eds.) *Eurosla Yearbook 1*, Benjamins, Amsterdam, 225-37.
- FIGUERAS, N., NORTH B., TAKALA S., VERHELST N. & VAN AVERMAET P. (2005), Relating examinations to the Common European Framework: A manual, in *Language Testing* 22, 261-79.
- FILLMORE C.J. (1979), On fluency, in C. Fillmore, D. Kempler & W.S.Y. Wang (eds.), *Individual differences in language ability and language behaviour*, Academic Press, New York, 85-101.
- FORD C., FOX B.A. & THOMPSON S.A. (2002) (eds.), *The language of turn and sequence*, Oxford University Press, Oxford.
- FOSTER P. (1996), Doing the task better: How planning time influences students' performance, in J. Willis and D. Willis (eds.), *Challenge and Change in Language Teaching*, Heinemann, London, 126-35.
- FOSTER P. & SKEHAN P. (1996), The influence of planning on performance in task-based learning, in *Studies in Second Language Acquisition* 18, 299-324.
- FOSTER P. & SKEHAN P. (1997), Modifying the task: the effects of surprise, time and planning type on task based foreign language instruction, in *Thames Valley University Working Papers in English Language Teaching* 4, 38-50.
- FOSTER P. & SKEHAN P. (1999), The influence of source of planning and focus of planning on task-based performance, in *Language Teaching Research* 3, 215-47.
- FOSTER P., TONKYN A. & WIGGLESWORTH G. (2000), Measuring spoken language: a unit for all reasons, in *Applied Linguistics* 21, 354-75.
- FREED B. (2000) Is fluency in the eyes (and ears) of the beholder?, in H. Riggensbach (ed.), *Perspectives on Fluency*, The University of Michigan Press, Ann Arbor, 243-65.
- FREEDLE R. & KOSTIN I. (1993), The prediction of TOEFL reading item difficulty: implications for construct validity, in *Language Testing* 10, 133-70.
- GARMAN M. (1990), *Psycholinguistics*, Cambridge University Press, Cambridge.
- GASS S.M., MACKEY A., FERNANDEZ M. & ALVAREZ-TORRES M. (1999), The effects of task repetition on linguistic output, in *Language Learning* 49, 549-80.

- GASS S.M. & MACKEY A. (2007), *Data Elicitation for Second and Foreign Language Research*, Lawrence Erlbaum Associates, Mahwah, NJ.
- GEDDES M. & STURTRIDGE G. (1978), *Listening links*, Heinemann, London.
- GIACALONE RAMAT A. (1999a), Le strategie di collegamento tra proposizioni nell'italiano di germanofoni, in N. Dittmar & A. Giacalone Ramat (eds.), *Grammatik und Diskurs / Grammatica e discorso*, Tübingen, Stauffenberg.
- GIACALONE RAMAT A. (1999b), Functional typology and strategies of clause connection in second-language acquisition, in *Linguistics* 37, 519-48.
- GILBERT R. (2007), The simultaneous manipulation along the planning time and +/- Here-and-Now dimensions: effects on oral L2 production, in P. Garcia Mayo (ed.), *Investigating tasks in formal language learning*, Multilingual Matters, Clevedon, 44-68.
- GILES H. (1977) (ed.), *Language, ethnicity and intergroup relations*, American Press, London-New York.
- GOLDSCHNEIDER J. & DEKEYSER R. (2001), Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A meta-analysis of multiple determinants, in *Language Learning* 51, 1-50.
- GRAFFI G. (1994), *Sintassi*, Il Mulino, Bologna.
- GRAFFI G. (2001), *200 Years of Syntax. A Critical Survey*, Benjamins, Amsterdam.
- GREGG K. (1990), The Variable Competence Model of second language acquisition and why it isn't, in *Applied Linguistics* 11, 364-83.
- GREGO BOLLI G. (2006a), Il Quadro e la valutazione, in M. Mezzadri (a cura di), *Integrazione linguistica in Europa*, UTET Università, Torino, 220-46.
- GREGO BOLLI G. (2006b), *Progetti europei. Nuove prospettive sulla scia del Quadro Comune Europeo di Riferimento*, <http://www.cvcl.it/canale.asp?id=198>
- HALE G. (1988), Student major field and text content: Interactive effects on reading comprehension in the Test of English as a Foreign Language, in *Language Testing* 5, 49-61.
- HALLIDAY M.A.K., *An Introduction to Functional Grammar*. 2nd edition, Edward Arnold, London, 1994.
- HARGREAVES A. (1992), Cultures of teaching: A focus for change, in A. Hargreaves & M. Fullan (eds.), *Understanding teacher development*, Teacher College Press, New York, 216-40.
- HARMER J. (1991), *The Practice of English Language Teaching*, Longman, London.
- HAWKEY R. & BARKER F. (2004), Developing a common scale for the assessment of writing, in *Assessing Writing* 9, 122-59.
- HENNING G. (1983), Oral proficiency testing: Comparative validities of interview, imitation, and completion methods, in *Language Learning* 33, 315-32.

- HENNING G. (1990), Priority issues in the assessment of second language abilities, in *Foreign Language Annals* 23, 379-84.
- HERDINA P. & JESSNER U. (2002), *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*, Multilingual Matters, Clevedon.
- HILL K. (1998), The effect of task-taker characteristics on reactions to and performance on an oral English proficiency test, in A.J. Kunnan (ed.), *Validation in language assessment*, Lawrence Erlbaum, Mahwah, 209-29.
- HOEKJE B. & LINNELL K. (1994), 'Authenticity' in language testing: Evaluating spoken language tests for international teaching assistants, in *Studies in Second Language Acquisition* 15, 461-93.
- HORGAN J. (1995), The new social Darwinist, in *Scientific America* 273, 174-81.
- HOUSE J. (1966), Development of pragmatic fluency in English as a foreign language, in *Studies in Second Language Acquisition* 18, 225-252.
- HOWARD M. (2004), On the interactional effect of linguistic constraints on interlanguage variation: The case of past marking, in *IRAL International Review of Applied Linguistics* 42, 319-34.
- HUEBNER T. (1985), System and variability in interlanguage syntax, in *Language Learning* 35, 141-63.
- HULSTIJN J.H. (1985), Testing second language proficiency with direct procedures. A response to Ingram, in K. Hyltenstam & M. Pienemann (eds.), *Modelling and Assessing Second Language Development*, Multilingual Matters, Clevedon, 277-82.
- HULSTIJN J. & SCHOONEN R. (2006), *Scientific report of the ESF sponsored exploratory workshop: Bridging the gap between research on second-language acquisition and research on language testing*, University of Amsterdam, Amsterdam Center for Language and Communication, [http://www.esf.org/fileadmin/be\\_user/ew\\_docs/05-208\\_Report.pdf](http://www.esf.org/fileadmin/be_user/ew_docs/05-208_Report.pdf)
- HUNT K.W. (1965), *Grammatical structures written at three grade levels*, The National Council of Teachers of English, Urbana IL.
- HUNT K.W. (1970), Recent measures in syntactic development, in M. Lester (ed.), *Reading in applied transformational grammar*, Holt. Rinehart, New York, 187-200.
- HYMES D. (1971), Competence and performance in linguistic theory, in R. Huxley & E. Ingram (eds.), *Language Acquisition, Models and Methods*, American Press, New York, 3-28.
- HYMES D. (1972), On communicative competence, in J.B. Pride & J. Holmes (eds.), *Sociolinguistics*, Penguin, Harmondsworth, 269-93.
- INGRAM D.E. (1985), Assessing proficiency: An overview of some aspects of testing, in K. Hyltenstam & M. Pienemann (eds.), *Modelling and assessing second language*

- acquisition*, Multilingual Matters, Clevedon, Avon, 215-76.
- ISHIKAWA T. (1995), Objective measurement of low-proficiency EFL narrative writing, in *Journal of Second Language Writing* 4, 51-70.
- ISHIKAWA T. (2007), The effect of increasing task complexity along the +/- Here-and-Now dimension, in P. Garcia Mayo (ed.), *Investigating tasks in formal language learning*, Multilingual Matters, Clevedon, 136-56.
- IWASHITA N., BROWN A., MCNAMARA T. & O'HAGAN S. (2007), Assessed levels of second language proficiency: How distinct?, in *Applied Linguistics Advance Access* published October, 26, 1-26.
- IWASHITA N., MCNAMARA T. & ELDER C. (2001), Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design, in *Language Learning* 51, 401-36.
- JAFRANCESCO E. (a cura di) (2004), *Le tendenze innovative del quadro comune europeo di riferimento per le lingue ed il portfolio. Atti del XII convegno nazionale ILSA*, Edilingua, Firenze.
- KAFTANDJIEVA F. & TAKALA S. (2002), Council of Europe scales of language proficiency: a validation study, in Alderson (ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case studies*, Council of Europe, Strasbourg, 106-29.
- KANE M. (1992), An argument-based approach to validity, in *Psychological Bulletin* 112, 527-35.
- KLEIN W. & DITTMAR N. (1979), *Developing Grammars: The Acquisition of German Syntax by Foreign Workers*, Springer, Berlin.
- KLEIN W. & PERDUE C. (1992), *Utterance structure. Developing Grammars Again*, John Benjamins, Amsterdam
- KUIKEN F. & VEDDER I. (2004a), Il Bed and Breakfast piu' bello d'Italia. Cognitive taakcomplexiteit en tekstkwaliteit in Italiaans T2 [Il Bed and Breakfast piu' bello d'Italia. Cognitive task complexity and text quality in Italian L2], in *Incontri* 19, 31-39.
- KUIKEN F. & VEDDER I. (2004b), De relatie tussen cognitieve taak complexiteit en linguïstische performance bij het schrijven in T1 en T2 [The relation between cognitive task complexity and linguistic performance in writing in L1 and L2], in *Toegepaste Taalwetenschap in Artikelen*, 72, 23-32.
- KUIKEN F. & VEDDER I. (2007a), Task complexity and linguistic complexity in L2 writing: a discussion, in *IRAL Special Issue on task complexity and instructed SLA*, 261-84.
- KUIKEN F. & VEDDER I. (2007b), Cognitive task complexity and linguistic performance in French L2 writing, in M.P. Garcia Mayo (ed.), *Investigating tasks in formal language learning*, Multilingual Matters, Clevedon, UK, 117-35.

- KUIKEN F. & VEDDER I. (2007c), Task complexity and measures of linguistic performance in L2 writing, in *IRAL* 45, 261-84.
- KUIKEN F. & VEDDER I. (2008), Cognitive task complexity and written output in Italian and French as a second language, in *Journal of Second Language Writing* 17, 48-60.
- KUIKEN F., MOS M. & VEDDER I. (2005), Cognitive task complexity and second language writing performance, in S. Foster-Cohen, M.P. Garcia-Mayo & J. Cenoz (eds.), *Eurosla yearbook. Vol. 5*, John Benjamins, Amsterdam, 195–222.
- KUNNAN A.J. (1993), Approaches to validation in language assessment, in A.J. Kunnan (ed.), *Validation in language assessment*, Lawrence Erlbaum, Mahwah, 1-18.
- LABOV W. (1970), The study of language in its social context, in *Studium Generale* 23, 30-87.
- LANGMAN J. & BAYLEY R. (2003), The Acquisition of Verbal Morphology by Chinese Learners of Hungarian, in *Language Variation and Change* 14, 55-77.
- LANGUAGE TESTING (2002), *Special issue on task-based language testing* 19.
- LANGUAGE TESTING (2005), *Special issue on the Common European Framework of Reference* 22.
- LANTOLF J.P. (2000), *Sociocultural Theory and Second Language Learning*, Oxford University Press, Oxford.
- LANTOLF J.P. & FRAWLEY W. (1985), Oral proficiency testing: a critical analysis, in *The Modern Language Journal* 69, 337-45.
- LANTOLF J.P. & FRAWLEY W. (1988), Proficiency: Understanding the construct, in *Studies in Second Language Acquisition* 10, 181-96.
- LANTOLF J.P. & FRAWLEY W. (1992), Rejecting the OPI – again: a response to Hagen, in *ADFL Bulletin* 23, 34-7.
- LARSEN-FREEMAN D. (1978), An ESL index of development, in *TESOL Quarterly* 12, 439-48.
- LARSEN-FREEMAN D. (1983), Assessing global second language proficiency, in H.W. Seliger & M. Long (eds.), *Classroom-oriented research in second language acquisition*, Newbury House, Rowley, MA, 287-304.
- LARSEN-FREEMAN D. (1997), Chaos/complexity science and second language acquisition, in *Applied Linguistics* 18, 141–65.
- LARSEN-FREEMAN D. (2002), Language acquisition and language use from a chaos/complexity theory perspective, in C. Kramsch (ed.), *Language Acquisition and Language Socialization*, Continuum, London.
- LARSEN-FREEMAN D. (2006), The emergence of complexity, fluency and accuracy in the oral and written production of five chinese learners of English, in *Applied Linguistics*, Special Issue 27, 590-619.

- LARSEN-FREEMAN D. (2007), On the complementarity of Chaos/Complexity Theory and Dynamic Systems Theory in understanding the second language acquisition process, in *Bilingualism: Language and Cognition* 10, 35-7.
- LARSEN-FREEMAN D. & CAMERON L. (2008), *Complex Systems and Applied Linguistics*, Oxford University Press, Oxford.
- LAUFER B. (2001), Quantitative evaluation of vocabulary, in in C. Elder, A. Brown, N. Iwashita, E. Grove, K. Hill, T. Lumley, T. McNamara & K. O'Loughlin (eds.), *Experimenting with uncertainty. Essays in honour of Alan Davies*, Cambridge University Press, Cambridge, 241-50.
- LAZARATON A. (1996), Interlocutor support in oral proficiency interviews: the case of CASE, in *Language Testing* 13, 151-72.
- LEHMANN C. (1988), Towards a typology of clause linkage, in J. Haiman & S.A. Thompson (eds.), *Clause combining in grammar and discourse*, Benjamins, Amsterdam.
- LENNON P. (1990), Investigating fluency in EFL: A quantitative approach, in *Language Learning* 40, 387-417.
- LEVELT W. (1989), *Speaking: From Intention to Articulation*, Cambridge University Press, Cambridge.
- LINACRE J.M. (1989b), *Many-faceted Rasch measurement*, MESA Press, Chicago.
- LIU G. (1991), *Interaction and second language acquisition: a case study of a Chinese child's acquisition of English as a second language*, Unpublished PhD thesis, La Trobe University, citato in E. Tarone & G. Liu (1995).
- LOBAN W. (1976), *Language development, Kindergarten through Grade 12*, Research Report No 18, NCTE Illinois – syntactic structures.
- LONG M. (1985a), A role for instruction in second language acquisition: Task-based language training, in K. Hyltenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition*, Multilingual Matters, Clevedon, Avon, 77-100.
- LONG M. (1985b), Input in second language acquisition theory, in S.M. Gass & C. Madden (eds.) *Input and Second Language Acquisition*, Newbury House, Rowley, MA, 377-93.
- LONG M. (1989), Task, group, and task.group interaction, in *University of Hawaii Working Papers in English as a Second Language* 8, 1-26.
- LONG M. & CROOKES G. (1992), Three approaches to task based syllabus design, in *TESOL Quarterly* 26, 27-56.
- LONG M. & ROBINSON P. (1998), Focus on form: theory, research and practise, in C. Doughty & J. Williams (eds.), *Focus on Form in Classroom SLA*, Cambridge University Press, Cambridge, 15-41.
- LONGACRE R. (1985), Sentences as combinations of clauses, in T. Shopen (ed.), *Language typology and syntactic description, vol II*, Cambridge University Press, Cambridge.

- LOSCHKY L. & BLEY-VROMAN R. (1993), Grammar and task-based methodology, in G. Crookes & S.M. Gass, *Tasks and Language Learning: Integrating Theory and Practice*, Multilingual Matters, Clevedon, Avon, 123-67.
- LUMLEY T. & MCNAMARA T. (1995), Rater characteristics and rater bias: implications for training, in *Language Testing* 12, 54-71.
- LUNZ M.E. & STAHAL J. (1990), Judge consistency and severity across grading periods, in *Evaluation and the Health Professions* 13, 425-44.
- LUNZ M.E., WRIGHT B.D. & LINACRE, J.M. (1990), Measuring the impact of judge severity on examination scores, in *Applied Measurement in Education* 3, 331-45.
- LYNCH T. & MCLEAN J. (2001), A case of exercising: effects of immediate task repetition on learners' performance, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*, Longman, London, 141-61.
- MACWHINNEY B. (1998), Models of emergence of language, in *Annual Review of Psychology* 49, 199-227.
- MACWHINNEY B. (1999) (ed.), *The emergence of language*, Lawrence Erlbaum, Mahwah, NJ.
- MACWHINNEY B. (2001), The gradual evolution of language, in B. Malle & T. Givón (eds.), *The evolution of language*, Benjamins, Philadelphia, 1-28.
- MACWHINNEY B. (2006), Emergentism – Use often and with care, in *Applied Linguistics*, Special Issue 27, 729-40.
- MAJOR R. (2004), Gender and stylist variation in second language phonology, in *Language Variation and Change* 16, 169-188.
- MCINTYRE P.N. (1993), *The importance and effectiveness of moderation training on the reliability of teacher assessments of ESL writing samples*, Unpublished MA Thesis, University of Melbourne, citata in McNamara (1996).
- MCLAUGHLIN B. (1990), Restructuring, in *Applied Linguistics* 11, 113-28.
- MCNAMARA T.F. (1997), Performance testing, in C. Clapham & D. Corson (eds.), *Language testing and assessment, Encyclopedia of language and education*, vol 7, Kluwer Academic, Dordrecht, 131-9.
- MCNAMARA T.F. (1995), Modelling performance: Opening Pandora's box, in *Applied Linguistics* 16, 159-79.
- MCNAMARA T.F. (1996), *Measuring Second Language Performance*, Longman, London.
- MCNAMARA T.F. (2006), Validity in language testing: The challenge of Sam Messick's legacy, in *Language Assessment Quarterly* 3, 31-51.
- MCNAMARA T.F., HILL K. & MAY L. (2002), Discourse and assessment, in *Annual Review of Applied Linguistics* 22, 221-42.



- MEHNERT U. (1998), The effects of different lengths of time for planning on second language performance, in *Studies in Second Language Acquisition* 20, 52-83.
- MEISEL J.M. (1997), The acquisition of the syntax of negation in French and German: Contrasting first and second language development, in *Second Language Research* 13, 227-63.
- MEISEL J.M., CLASHEN H. & PIENEMANN M. (1981), On determining developmental stages in second language acquisition, in *Studies in Second Language Acquisition* 3, 109-35.
- MESSICK S. (1989), Validity, in R.L. Linn (ed.), *Educational measurement*, American Council in Education and Macmillan, 13-103.
- MESSICK S. (1996), Validity and washback in language testing, in *Language Testing* 13, 241-56.
- MEZZADRI M. (2004), *Il Quadro comune europeo a disposizione della classe. Un percorso verso l'eccellenza*. Guerra, Perugia.
- MICHEL M.C., KUIKEN F. & VEDDER I. (2007), The influence of complexity in monologic versus dialogic tasks in Dutch L2, in *IRAL, International Review of Applied Linguistics* 45, 241-59.
- MILLER J. & WEINERT R. (1998), *Spontaneous spoken language: syntax and discourse*, Clarendon, Oxford.
- MINISTERO DELL'ISTRUZIONE, DELL'UNIVERSITÀ E DELLA RICERCA (2008), *Gli alunni stranieri nel sistema scolastico italiano, anno scolastico 2007/2008*, in [www.istruzione.it](http://www.istruzione.it).
- MINISTERO DELLA PUBBLICA ISTRUZIONE (2003), *Linee guida per l'integrazione*, in [www.istruzione.it](http://www.istruzione.it)
- MINISTERO DELLA PUBBLICA ISTRUZIONE (2006), *Alunni con cittadinanza non italiana. Scuole statali e non statali, anno scolastico 2005/2006*, in [www.istruzione.it](http://www.istruzione.it).
- MISLEVY R.J., STEINBERG L.S. & ALMOND R.G. (2003), On the structure of assessment arguments, in *Measurement: Interdisciplinary Research and Perspective* 1, 3-62.
- MIYAKE A. & SHAH P. (1999), *Models of working memory: Mechanism of active maintenance and executive control*, Cambridge University Press, Cambridge.
- MÖHLE D. (1984), A comparison of the second language production of different native speakers, in H.W. Dechert, D. Möhle & Raupach M. (eds.), *Second Language Productions*, Gunter Narr, Tübingen, 26-49.
- MONEGLIA M. (1999), *A note on spoken language corpora, units of analysis and language sampling strategies*, paper presented at the workshop Morphosyntax (text corpora and tagging), Scientific Network Intersign: Sign Linguistics and data Exchange, European Science Foundation, Certosa di Pontignano, March, 12-15, in <http://www.sign-lang.uni-hamburg.de/intersign/internal/>.
- MONEGLIA M. (2000a), Corpus LABLITA, corpus italien de langue parlée, in M. Bilger (ed.),

- Corpus. Méthodologie et applications linguistiques*, Honoré Champion, Paris.
- MONEGLIA M. (2000b), Units of analysis of spontaneous speech and speech variation in a cross-linguistic perspective, in <http://lablita.dit.unifi.it/italian/preprint.html>
- MONEGLIA M. (2004), Measurements of Spoken Language Variability in a Multilingual Corpus. Predictable aspects, in M.T. Lino, M.F. Xavier, F. Ferreira, R. Costa & R. Silva (eds.), *Proceeding of the 4th LREC Conference*, ELRA, Paris, vol. 4, 1419-22.
- MONEGLIA M. & CRESTI E. (1997), L'intonazione e i criteri di trascrizione del parlato, in U. Bortolini & E. Pizzuto (a cura di), *Il progetto CHILDES Italia*, vol. II, Il Cerro, Pisa, 57-90.
- MONROE J.H. (1975), Measuring and enhancing syntactic fluency in French, in *The French Review* 48, 1023-31.
- MORROW K. (1979-81), Communicative language testing: Revolution or evolution?, in C.J. Brumfit & K. Johnson (eds.), *The communicative approach to language teaching*, Oxford University Press, Oxford, reprinted in J.C. Alderson & A. Huges, (eds.) (1981), *Issues in language testing*, ELT Document 111, British Council, London, 9-25.
- MOUGEON R. & REHNER K. (2001), Acquisition of sociolinguistic variants by French immersion students: The case of restrictive expressions, and more, in *Modern Language Journal* 85, 398-415.
- NEFF J., DAFOUZ E., DIEZ M., PRIETO R. & CHAUDRON C. (1998), *Contrastive discourse analysis: Argumentative text in English and Spanish*, Paper presented at the Twenty-Fourth Linguistics Symposium: Discourse across Languages and Cultures, University of Wisconsin-Milwaukee, September, citato in Ortega (2003).
- NEMETH N. & KORMOS J. (2001), Pragmatic aspects of task performance: the case of argumentation, in *Language Teaching Research* 5, 213-40.
- NEWELL A. & SIMON H. (1972), *Human Problem Solving*, Prentice Hall, Englewood Cliffs, NJ.
- NIWA (2000), citato in P. Robinson (2001a) (ed.), *Cognition and second language instruction*, Cambridge University Press, Cambridge.
- NORRIS J., BROWN J.D., HUDSON T. & YOSHIOKA J. (1998), *Designing second language performance assessments*, University of Hawaii at Manoa, Honolulu.
- NORRIS J. & ORTEGA L. (2003), Defining and measuring SLA, in C.J. Doughty & M.H. Long (eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, Malden, MA, 717-61.
- NORRIS J. & ORTEGA L. (in press), Measurement for understanding: An organic approach to investigating complexity, accuracy, and fluency in SLA, in A. Housen & F. Kuiken (eds.), *Fluency, accuracy and complexity in second language acquisition: Theoretical and methodological perspectives*, in *Applied Linguistics* Special Issue 30.

- NORTH B. (1991), Standardisation of continuous assessment grades, in C.J. Alderson & B. North (eds.), *Language Testing in the 1990s*, Macmillan, London, 167-77.
- NORTH B. (2000), *The development of a common european framework scale of language proficiency*, Peter Lang, New York.
- NORTH B. (2002), A CEF-based self assessment tool for university entrance, in J.C. Alderson (ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case studies*, Council of Europe, Strasbourg, 146-66.
- NORTH B. & SCHNEIDER G. (1998), Scaling descriptors for language proficiency scales, in *Language Testing* 15, 217-62.
- NUNAN D. (1989), *Designing Tasks for the Communicative Classroom*, Cambridge University Press, Cambridge.
- OCHS E., SCHEGLOFF E. & THOMPSON S.A. (1996) (eds.), *Interaction and Grammar*, Cambridge University Press, Cambridge.
- OLLER J.W. (1979), *Language tests at school. A pragmatic approach*, Longman, London.
- OLLER J.W. (1983), A consensus for the Eighties?, in J.W. Oller (ed.), *Issues in language testing research*, Newbury House, Rowley (Mass.), 351-56.
- ORTEGA L. (1999), Planning and focus on form in L2 oral performance, in *Studies in Second Language Acquisition* 21, 108-48.
- ORTEGA L. (2003), Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of College-level L2 writing, in *Applied Linguistics* 24, 492-518.
- OSCARSON M. (1978), *Approaches to Self-assessment in Foreign Language Learning*, Council of Europe, Strasbourg.
- OSCARSON M. (1984), *Self-assessment of Foreign Language Skills: a Survey of Research and Development Work*, Council of Europe, Strasbourg.
- OSSERVATORIO NAZIONALE (2007), *La via italiana per la scuola interculturale e l'integrazione degli alunni stranieri*, in [www.istruzione.it](http://www.istruzione.it).
- PALLOTTI G. (2007), An operational definition of the emergence criterion, in *Applied Linguistics* 28, 361-82.
- PALLOTTI & FERRARI (2008), La variabilità situazionale dell'interlingua: implicazioni per la ricerca acquisizionale e il testing, in G. Bernini, L. Spreafico e A. Valentini (a cura di), *Competenze lessicali e discorsive nell'acquisizione di lingue seconde*, Perugia, Guerra, 437-462.
- PERKINS K., GUPTA L. & TAMMANA R. (1995), Predicting item difficulty in a reading comprehension test with an artificial neural network, in *Language Testing* 12, 34-53.
- PIENEMANN M. (1985), Learnability and syllabus construction, in K. Hyttenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition*, Multilingual Matters, Clevedon, Avon, 23-75.

- PIENEMANN M. (1987), Determining the influence of instruction on L2 speech processing, in *Australian Review of Applied Linguistics* 10, 83-113.
- PIENEMANN M. (1988), *Language Processing and Second Language Development: Processability Theory*, Benjamins, Amsterdam.
- PIENEMANN M. (1992), COALA: a computational system for interlanguage analysis, in *Second Language Research* 8, 59-92
- PIENEMANN M. (1998), *Language Processing and Second-Language Development: Processability Theory*, John Benjamins, Amsterdam.
- PIENEMANN M. (2007), Variation and dynamic systems in SLA, in *Bilingualism: Language and Cognition* 10, 43-5.
- PIENEMANN M. (2008), A brief introduction to processability theory, in J.-U. Keßler (ed.), *Processability Approaches to Second Language Development and Second Language Learning*, Cambridge University Press, Cambridge, 9-30.
- PIENEMANN M. & MACKEY A. (1992), An empirical study of children ESL development and Rapid Profile, in P. McKay (ed.), *ESL development: Language and Literacy in Schools project. Vol. II*. National Languages and Literacy Institute of Australia, 115-259.
- POLIO C.G. (1997), Measures of linguistic accuracy in second language writing research, in *Language Learning* 47, 101-43
- PRABHU N.S. (1987), *Second Language Pedagogy*, Oxford University Press, Oxford.
- PRESTON D. (1989), *Sociolinguistics and Second Language Acquisition*, Blackwell, Oxford.
- PRESTON D. (1993), Variation linguistics and SLA, in *Second Language Research* 9, 153-72.
- PRESTON D. (1996), Variationist linguistics and second language acquisition, in W.C. Ritchie & T.K. Bhatia (eds.), *Handbook of Second Language Acquisition*, Academic Press, New York, 229-65.
- PRESTON D. (2000), Three kinds of sociolinguistics and SLA: a psycholinguistics perspective, in B. Schwierzbin, F. Morris, M. Anderson, C. Klee & E. Tarone (eds.), *Social and Cognitive Factors in Second Language Acquisition*, Cascadilla Press, Somerville, MA, 3-30.
- PURPURA J.E. (1993), The development and construct validation of an instrument designed to investigate selected cognitive background characteristics of test-takers, in A.J. Kunnan (ed.), *Validation in language assessment*, Lawrence Erlbaum, Mahwah, 114-40.
- PURPURA J.E. (1997), An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance, in *Language Learning* 47, 289-325.
- PURPURA J.E. (1999), *Modelling the relationships between test takers' reported cognitive and metacognitive strategy use and performance on language tests*, in University of Cambridge Local Examinations Syndacate and Cambridge University Press,

Cambridge.

- RAHIMPOUR M. (1997), *Task condition, task complexity and variation in oral discourse*, unpublished PHD dissertation, citato in P. Robinson (2001).
- RAHIMPOUR M. (1999), Task complexity and variation in interlanguage, in N. Jungheim & P. Robinson (eds.), *Pragmatics and Pedagogy: Proceedings of The 3rd Pacific Second Language Research Forum*, Vol2, PacSLRF, Tokyo, 115-34.
- RAUPACH M. (1980), Temporal variables in first and second language production, in H.W. Dechert & M. Raupach (eds.), *Temporal Variables in Speech: Studies in Honor of Freida Goldman Eissler*, Mouton, The Hague.
- READ J. (2007), Towards a new collaboration: research in SLA and language testing, in *New Zealand Studies in Applied Linguistics* 13, 22-35.
- REGAN V. (1995), The acquisition of sociolinguistic native speech norms: Effects of a year abroad on second language learners of French, in B. Freed, (ed.), *Second Language Acquisition in a Study Abroad Context*, Benjamins, Amsterdam and Philadelphia, 245-67.
- REGAN V. (1996), Variation in French interlanguage: A longitudinal study of sociolinguistic competence, in R. Bayley & D. Preston (eds.), *Second Language Acquisition and Linguistic Variation*, John Benjamins, Amsterdam, 177-201.
- REGAN V. (1997), Les apprenants avancés, la lexicalisation et l'acquisition de la compétence sociolinguistique: une approche variationiste, in *AILE: Acquisition et Interaction en Langue Etrangère* 9, 193-210.
- REGAN V. (2004), The relationship between the group and the individual and the acquisition of native speaker variation patterns: A preliminary study, *IRAL International Review of Applied Linguistics* 42, 335-48.
- REHNER K. & MOUGEON R. (1999), Variation in the spoken French of immersion students: to *ne* or not to *ne*, that is the sociolinguistic question, in *Canadian Modern Language Review* 56, 124-56.
- REHNER K., MOUGEON R. & NADASDI T. (2003), The learning of sociolinguistic variation by advanced FLS learners: The case of *nous* versus *on* in immersion French, in *Studies in Second Language Acquisition* 25, 127-56.
- RENZI L. & SALVI G. (2001) (a cura di), *Grande grammatica italiana di consultazione*, vol. II. *I sintagmi verbale, aggettivale, avverbale. La subordinazione*, Il Mulino, Bologna.
- RENZI L., SALVI G. & CARDINALETTI A. (1988-1995), *Grande grammatica italiana di consultazione*, vol I°-III°, Il Mulino, Bologna.
- RIMMER W. (2006), Measuring grammatical complexity: The Gordian knot, in *Language Testing* 23, 497-519.
- ROBINSON P. (1995), Task complexity and second language narrative discourse, in *Language*

*Learning* 45, 99-140.

- ROBINSON P. (1996), Attention and memory during SLA, in C.J. Doughty & M.H. Long (eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, 631-78.
- ROBINSON P. (2000), *Task complexity and reasoning demands: effects on dyadic NNS-NNS interaction, fluency, accuracy and incorporation of input*. Unpublished data, Aoyama Gakuin University, Tokyo, citato in Robinson (2001a).
- ROBINSON P. (2001a) (ed.), *Cognition and second language instruction*, Cambridge University Press, Cambridge.
- ROBINSON P. (2001b), Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA, in P. Robinson (ed.), *Cognition and Second Language Instruction*, Cambridge University Press, Cambridge, 287-318.
- ROBINSON P. (2001c), Task complexity, task difficulty, and task production: Exploring interaction in a componential framework, in *Applied Linguistics* 22, 27-57.
- ROBINSON P. (2003a), Attention and memory, in C.J. Doughty & M.H. Long (eds.), *The Handbook of Second Language Acquisition*, Blackwell, Oxford, 631-78.
- ROBINSON P. (2003b), The Cognition Hypothesis, task design and adult task-based second language learning, in *Second Language Studies*, 21, 45-105.
- ROBINSON P. (2005a), Cognitive complexity and task sequencing: A review of studies in a componential framework for second language task design, in *IRAL Internationa Review of Applied Linguistics* 43, 1-33.
- ROBINSON P. (2005b), Aptitude and second language acquisition, in *Annual Review of Applied Linguistics* 25, 45-73.
- ROBINSON P. (2007a), Criteria for classifying and sequencing pedagogic tasks, in M. García Mayo (ed.), *Investigating tasks in formal language settings*, Multilingual Matters, Clevedon, 7-25.
- ROBINSON P. (2007b), Task complexity, theory of mind, and intentional reasoning: effects of L2 speech production, interaction, uptake and perceptions of task difficulty, in *IRAL Internationa Review of Applied Linguistics* 45, 193-213.
- ROBINSON P. (2007c), Aptitude, abilities, contexts, practice, in R. DeKeyser (ed.), *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology*, Cambridge University Press, New York, 256-286.
- ROBINSON P. & ELLIS N.C. (eds.) (2008), *Handbook of Cognitive Linguistics and Second Language Acquisition*, Routledge, London.
- ROMAINE S. (1996), Variation, in C.J. Doughty & M.H. Long (eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, London, 409-35.
- SAMUDA V. (2001), Guiding relationship between form and meanng during task performance: the role of the teacher, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching*

- Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, 119-40.
- SANKOFF D. (1998), Variable rules, in U. Ammon, N. Dittmar & K.J. Mattheier (eds.), *Sociolinguistics: An International Handbook of Science of Language and Society* vol2, Mouton de Gruyter, Berlin, 984-97.
- SANKOFF D. & LABOV W. (1979), On the use of variable rules, in *Language in Society* 8, 189-222.
- SANKOFF D., THIBAUT P., NAGY N., BLONDEAU E., FONOLOSSA M. & GAGNON L. (1997), Variation in the use of discourse markers in contact situation, in *Language Variation and Change* 9, 191-217.
- SASAKI M. (1996), *Second language proficiency, foreign language aptitude, and intelligence: quantitative and qualitative analyses*, Peter Lang, New York.
- SCARANO A. (1996), Frasi relative e pseudorelative in italiano, in *Studi di grammatica italiana* XVI, 377-423.
- SCARANO A. (2003), Relative appositive e aggettivi appositivi. Tra sintassi e articolazione dell'informazione, in P. D'Achille (a cura di), *Atti del convegno internazionale della SILFI, Generi, architetture e forme testuali*, Roma 1-5 ottobre 2002, in preprint LABLITA 1, 1-12, in <http://lablita.dit.unifi.it/italian/preprint.html>
- SCHEGLOFF E.A. (1979), Identification and recognition in telephone conversation openings, in G. Psathas (ed.), *Everyday Language: studies in ethnomethodology / conversation analysis*, Irvington Publishers, New York.
- SCHEGLOFF E.A. (1986), The Routine as Achievement, in *Human Studies* 9, 111-51.
- SCHMIDT R.W. (1992), Psychological mechanism underlying second language fluency, in *Studies in Second Language Acquisition* 14, 357-85.
- SCHNEIDER G. & NORTH B. (2000), Fremdsprachen Können: was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit, in *Nationales Forschungsprogramm 33: Wirksamkeit unserer Bildungssysteme*, Verlag Rüegger, Chur, Zürich.
- SCHOONEN R., VERGEER M. & EITING M. (1997), The assessment of writing ability: expert readers versus lay readers, in *Language Testing* 14, 157-84.
- SCOTT C.M. (1988), Spoken and written syntax, in M.A. Nippold (ed.), *Lateral language development. Ages 9 through 19*, College Hill Press, Boston, 49-95.
- SELINKER L. (1969), Language transfer, in *General Linguistics* 9, 67-92.
- SELINKER L. (1972), Interlanguage, in *IRAL International Review of Applied Linguistics* 10, 219-31.
- SELINKER L. (1981), Updating the interlanguage hypothesis, in *Studies in Second Language Acquisition* 3, 201-20.

- SERIANNI L. (1988), *Grammatica italiana. Italiano comune e lingua letteraria*, Torino, UTET.
- SHARMA A. (1980), Syntactic maturity: Assessing writing proficiency in a second language, in R. Silverstein (ed.), *Occasional Papers in Linguistics* 6, Southern Illinois University, Carbondale.
- SHOHAMY E. (1983), The stability of the oral proficiency trait on the Oral Interview Speaking Test, in *Language Learning* 33, 527-40.
- SHOHAMY E. (1994), The role of language tests in the construction and validation of second language acquisition theories, in E. Tarone, S.M. Gass & A.D. Cohen (eds.), *Research methodology in second language acquisition*, Lawrence Erlbaum, Hillsdale, NJ, 115-31.
- SHOHAMY E. (1998), How can language testing and SLA benefit from each other? The case of discourse, in L.F. Bachman & A.D. Cohen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, 156-76.
- SHOHAMY E. (2000), The relationship between language testing and second language acquisition, revisited, in *System* 28, 541-53.
- SHOHAMY E., GORDON C.M. & KRAEMER R. (1992), The effect of raters' background and training on the reliability of direct writing tests, in *Modern Language Journal* 76, 27-33.
- SHOHAMY E., REVES T. & BEJARANO Y. (1986), Introducing a new comprehensive test of oral proficiency, in *English Language Teaching Journal* 40, 212-22.
- SIMONE R. (1990), *Fondamenti di linguistica*, Bari, Laterza.
- SKEHAN P. (1989), *Individual Differences in Second-Language Learning*, Edward Arnold, London.
- SKEHAN P. (1995a), A framework for the implementation of task-based instruction, in *Applied Linguistics* 16, 542-566.
- SKEHAN P. (1995b), Analysability, accessibility and ability for use, in G. Cook & S. Seidlhofer (eds.), *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson*, Oxford University Press, Oxford, 91-106.
- SKEHAN P. (1996a), A framework for the implementation of task-based instruction, in *Applied Linguistics* 17, 38-62.
- SKEHAN P. (1996b), Second language acquisition research and task-based instruction, in J. Willis & D. Willis, (eds.), *Challenge and Change in Language Teaching*, Heinemann, Oxford, 17-30.
- SKEHAN P. (1998), *A cognitive approach to language learning*, Cambridge University Press, Cambridge.
- SKEHAN P. (2001), Task and language performance assessment, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, 167-85.



- SKEHAN P. (2001), Task and language performance assessment, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*, Pearson Education, London, 167-85.
- SKEHAN P. (2003), Task-based instruction, in *Language Teaching* 36, 1-14.
- SKEHAN P. & FOSTER P. (1997), Task type and task processing conditions as influence on foreign language performance, in *Language Teaching Research* 1, 1-27.
- SKEHAN P. & FOSTER P. (1999), The influence of task structure and processing conditions on narrative retellings, in *Language Learning* 49, 93-120.
- SKEHAN P. & FOSTER P. (2001), Cognition and tasks, in P. Robinson (ed.), *Cognition and second language instruction*, Cambridge University Press, Cambridge, 183-205.
- SKEHAN P. & FOSTER P. (2005), Strategic on-line planning: The influence of surprise information and task time on second language performance, in R. Ellis (ed.), *Planning and Task Performance in a Second Language*, John Benjamins, Amsterdam/Philadelphia, 183-205.
- SLAVOFF G.R. & JOHNSON J.S. (1995), The effects of age on the rate of learning a second language, in *Studies in Second Language Acquisition* 17, 1-16.
- SMITH J. (1989), Topic and variation in ITA oral proficiency: SPEAK and field specific test, in *English for Specific Purposes* 8, 155-67.
- SORNICOLA R. (1981), *Sul parlato*, Il Mulino, Bologna.
- SPARKS R.L., ARTZER M., GANSCHOW L., SIEBENHAR D., PLAGEMAN M. & PATTON J. (1998), Differences in native-like language skills, foreign language aptitude, and foreign language grades among high-, average-, and low-proficiency foreign-language learners: two studies, in *Language Testing* 15, 181-216.
- STEPHAN A. (1999), Varieties of emergentism, in *Evolution and Cognition* 5, 49-59.
- STERN H.H. (1983), *Fundamental concepts of language teaching*, Oxford University Press, New York.
- SWAIN M. (1998), Focus on form through conscious reflection, in G. Doughty & J. Williamson (eds.), *Focus on Form in Classroom Second Language Acquisition*, Cambridge University Press, Cambridge, 64-81.
- SWAIN M. & LAPKIN S. (2001), Focus on Form through collaborative dialogue: exploring task effects, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, 99-118.
- TALMY L. (1978), Relations between subordination and coordination, in J. Greenberg (ed.) *Universals of human language, vol IV: Syntax*, Stanford University Press, Stanford, 487-513.
- TARONE E. (1979), Interlanguage as a camaleon, in *Language Learning* 29, 181-91.
- TARONE E. (1982), Systematicity and attention in interlanguage, in *Language Learning* 32, 69-

- TARONE E. (1983), On the variability of interlanguage systems, in *Applied Linguistics* 4, 143-63.
- TARONE E. (1985), Variability in Interlanguage use: A study of style-shifting in morphology and syntax, in *Language Learning* 35, 373-403.
- TARONE E. (1988), *Variation in Interlanguage*, Edward Arnold, London.
- TARONE E. (1990), On variation in interlanguage: A response to Gregg, in *Applied Linguistics* 11, 392-400.
- TARONE E. & LIU G. (1995), Situational context, variation, and second language acquisition theory, in G. Cook & B. Seidlhofer (eds.), *Principle and practice in applied linguistics. Studies in honour of H. G. Widdowson*, Oxford University Press, Oxford, 107-24.
- TARONE E. & PARRISH B. (1988b), Task-related variation in interlanguage: The case of articles, in *Language Learning* 38, 21-44.
- TAVAKOLI P. & SKEHAN P. (2005), Strategic planning, task structure, and performance testing, in R. Ellis (ed.), *Planning and task performance in a second language*, Benjamins, Amsterdam, 239-73.
- TAYLOR D.S. (1988), The meaning and use of the term 'competence' in linguistics and applied linguistics, in *Applied Linguistics* 9, 148-68.
- THELEN E. & BATES, E. (2003), Connectionism and dynamic systems: Are they really different?, in *Developmental Science* 6, 378-91.
- THELEN E. & SMITH L.B. (1994), *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press, Cambridge MA.
- THOMAS M. (1994), Assessment of L2 proficiency in second language acquisition, in *Language Learning* 44, 307-36.
- TOWELL R., HAWKINS R. & BAZERGUI N. (1996), The development of fluency in advanced learners of French, in *Applied Linguistics* 17, 84-119.
- TRIM J. (1978), *Some Possible Lines of Development of an Overall Structure for a European Unit/Credit Scheme for Foreign Language Learning by Adults*, Council of Europe, Strasbourg.
- VALENTINI A. (1988), Le frasi causali e l'emergere della subordinazione in italiano L2: il caso di due apprendenti cinesi, in *Linguistica e Filologia* 8, 113-48.
- VALENTINI A. (2001), La frase finale in italiano L2, in *Vox Romanica* 60, 69-88.
- VAN DAELE S., HOUSEN A. & PIERRARD M. (2007), Fluency, accuracy and complexity in the manifestation and development of two second languages, in S. Van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder, *Complexity, Accuracy and Fluency in second language use, learning and teaching*, Contactforum, Brussel, 301-16.
- VAN EK J. (1975), *The Threshold level in a European Unit/Credit System for Modern*

- Language Learning by Adults*, Council of Europe, Strasbourg, republished in 1977 as *The Threshold level for Modern Language Learning*, Longman, London.
- VAN EK J. & TRIM J.L.M. (1984) (eds.), *Across the Threshold*, Pergamon Press, Oxford.
- VAN EK J. & TRIM J.L.M. (1990), *Waystage 1990*, Cambridge University Press, Cambridge.
- VAN EK J. & TRIM J.L.M. (1991), *Threshold 1990, a Revisited and Extended Version of "The Threshold level" by J.A. van Ek*, Council of Europe, Strasbourg.
- VAN EK J. & TRIM J.L.M. (1995), *Vantage Level*, Council of Europe, Strasbourg.
- VAN GEERT P. & VAN DIJK M. (2002), Focus on variability: New tools to study intra-individual variability in developmental data, in *Infant Behaviour & Development* 25, 340-74.
- VAN GEERT P. (1991), A dynamic system mode of cognitive and language growth, in *Psychological Review* 98, 3-53.
- VAN GEERT P. (1994), Vygotskian Dynamics of Development, in *Human Development* 37, 346-65.
- VAN GEERT P. & STEENBEEK H. (2005), Explaining after by before: Basic aspects of a dynamic systems approach to the study of development, in *Developmental Review* 25, 408-42.
- VAN LIER L. (1989), Reeling, writhing, drawling, stretching, and fainting in coil: Oral proficiency interviews as conversation, in *TESOL Quarterly* 23, 489-508.
- VAN PATTEN B. (1990), Attending to content and form in the input: an experiment in consciousness, in *Studies in Second Language Acquisition* 12, 287-301.
- VAN PATTEN B. (1996), *Input Processing and Grammar Instruction*, Ablex, New York.
- VEDOVELLI M. (2002), *Guida all'italiano per stranieri: la prospettiva del quadro comune europeo*, Carocci, Roma.
- VOGHERA M. (1992), *Sintassi e intonazione dell'italiano parlato*, Il Mulino, Bologna.
- VOLLMER H.J. (1981), Why are we interested in 'general language proficiency', in J.C. Alderson & A. Huges (eds.), *Issues in language testing*, ELT Document 111, British Council, London, 152-76.
- VOLLMER H.J. & SANG F. (1983), Competing hypothesis about second language ability: A plea for caution, in J.W. Oller (ed.), *Issues in language testing research*, Newbury House, Rowley (Mass.), 29-79.
- WEIGLE S.C. (1994), *Effects of training on raters of ESL compositions: quantitative and qualitative approaches*, unpublished PhD dissertation, University of California, Los Angeles, citata in McNamara (1996).
- WEIGLE S.C. (1998), Using FACETS to model rater training effects, in *Language Testing* 15, 263-67.
- WEIR C.J. (2005), Limitations of the Common European Framework for developing comparable examinations and tests, in *Language Testing* 22, 281-300.

- WENDEL J.N. (1997), *Planning and second language narrative production*, unpublished doctoral dissertation, temple University, Japan, citato in Ellis (2003).
- WESCHE M. & SKEHAN P. (2002), Communicative teaching, content-based instruction, and task-based learning, in R. Kaplan (ed.), *Handbook of Applied Linguistics*, Oxford University Press, Oxford, 207-228.
- WIDDOWSON H.G. (1989), Knowledge of language and ability of use, in *Applied Linguistics* 10, 128-37.
- WIEMANN J.M. & BACKLUND P. (1980), Current theory and research in communicative competence, in *Review of Educational Research* 50, 185-99.
- WIESE R. (1984), Language production in foreign and native languages. Same or different?, in H. Dechert, D. Möhle & Raupach M. (eds.), *Second Language Productions*, Gunter Narr, Tübingen, 11-25.
- WIGGLESWORTH G. (1993) Exploring bias analysis as a tool for improving rater consistency in assessing oral interaction, in *Language Testing* 10, 305-35.
- WIGGLESWORTH G. (1997), An investigation of planning time and proficiency level on oral test discourse, in *Language Testing* 14, 85-106.
- WIGGLESWORTH G. (2001), Influences on performance in task-based oral assessments, in M. Bygate, P. Skehan & M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, 186-209.
- WILLIS J. (1996), *A Framework for Task-based learning*, Longman, London.
- WILLIS J. & WILLIS D. (1988), *COBUILD Book 1*, Collins, London.
- WOLFE-QUINTERO K., INAGAKI S. & KIM H.Y. (1998), *Second Language Development in writing: measures of fluency, accuracy, and complexity*, University of Hawaii, Second Language Teaching and Curriculum Center, Honolulu.
- WOLFRAM W. (1985), Variation in tense marking: A case for the obvious, in *Language Learning* 35, 229-53.
- WU Y. (1998), What do tests of listening comprehension test? A retrospection study of EFL test-takers performing a multiple-choice task, in *Language Testing* 15, 21-44.
- YOUNG R. (1988), Variation and the interlanguage hypothesis, in *Studies in Second Language Acquisition* 10, 281-302.
- YOUNG R. (1991), *Variation in Interlanguage Morphology*, Peter Lang, New York.
- YOUNG R. (1996), Form-function relations in articles in English interlanguage, in R. Bayley & D. Preston (eds.), *Second Language Acquisition and Linguistic Variation*, Benjamins, Amsterdam, 135-75.
- YOUNG R. (1999), Sociolinguistic approaches to SLA, in *Annual Review of Applied Linguistics* 19, 105-32.

- YUAN F. & ELLIS R. (2003), The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production, in *Applied Linguistics* 24, 1-27.
- ZUENGLER J. (1989), Assessing an interactive paradigm: how accomodative should we be? In M. Eisenstein (ed.), *The Dynamic Interlanguage: Empirical studies in second language variation*, Plenum Press, New York.



**Tabella 1. Task-based studies, produzione orale e CAF, pianificazione**

studio	soggetti	domande di ricerca	metodologia	variabili dipendenti	risultati
Foster (1996)	32 apprendenti di inglese L2 a livello intermedio	quali sono gli effetti della pianificazione su complessità, accuratezza e fluenza?	3 task: (1) informazioni personali (2) storia (3) decision making. Analisi dei primi 5 min  3 gruppi: (1) pianificazione guidata, (2) pianificazione libera (3) controllo pianificazione, 10 min	<i>complessità:</i> - varietà delle forme per il passato <i>clausole per c unit</i> <i>fluenza:</i> - numero di pause - silenzio totale - ripetizioni	<i>complessità:</i> la pianificazione dà una maggior varietà di forme passate e c unit di più clausole, soprattutto se guidata e nei task cognitivamente più complessi. <i>accuratezza:</i> la pianificazione non guidata dà maggior accuratezza in due task su tre, informazioni personali e decision making. <i>fluenza:</i> la pianificazione dà maggiore fluenza.
Foster & Skehan (1996)	come in Foster (1996)	ipotesi generale: la pianificazione ha un'influenza su complessità, accuratezza e fluenza. L'effetto del pianificazione dipenderà dal task.	come in Foster (1996)	<i>complessità:</i> - clausole/ c-units - varietà di forme verbali <i>accuratezza:</i> - clausole senza errori - errori lessicali <i>fluenza:</i> - riformulazioni - sostituzioni - false partenze - ripetizioni - esitazioni - pause (+ di 1 secondo) - silenzio totale	<i>complessità:</i> la pianificazione aumenta l'uso della subordinazione, soprattutto se guidata. La pianificazione dà una maggior varietà di forme verbali nei task (1) e (2), ma non in (3). <i>accuratezza:</i> la pianificazione dà maggior accuratezza nei task (1) e (3), ma non nel task (2). <i>fluenza:</i> per i task (1) e (3) la pianificazione dà più fluenti. Nessun effetto per tipo di pianificazione.

\*studi in cui si confrontano misure CAF e valutazioni soggettive

*Wigglesworth (1997)	107 adulti apprendenti di inglese L2 in Australia. 28 distinti in livello elementare e intermedio-avanzato.	che effetto ha la pianificazione su valutazioni e misure di complessità, accuratezza e fluenza?	5 task con un diverso grado di complessità cognitiva: - descrizione - differenze - storia - riassunto conversazione, - messaggio segreteria telefonica, - spiegazione - discussione  pianificazione, 1 min.	<i>valutazioni soggettive</i> grammatica, fluenza e comprensibilità.  <i>misure analitiche</i> <i>complessità:</i> - grado di subordinazione <i>accuratezza:</i> - uso della marca –s per il plurale - morfologia del verbo - articoli <i>fluenza:</i> - numero dei self-repair	la pianificazione dà maggior complessità, accuratezza e fluenza, soprattutto per gli apprendenti più avanzati e per i task cognitivamente più complessi. Tale tendenza non è confermata però per tutte le misure e tutti i task.
Skehan and Foster (1997)	40 apprendenti di inglese L2 a livello intermedio	che effetto ha la pianificazione sulla performance?	come in Foster (1996).  2 variabili: - pianificazione, 10 min - post task	<i>complessità:</i> - clausole per c-unit <i>accuratezza:</i> - clausole senza errori <i>fluenza:</i> - numero delle pause di durata superiore ad 1 secondo.	<i>complessità:</i> la pianificazione dà maggior complessità nei task (1) e (3), ma non in (2). <i>accuratezza:</i> la pianificazione dà maggior accuratezza solo nel task (2). <i>fluenza:</i> la pianificazione dà maggior fluenza nei task (1) e (2)  il post task genera maggior fluenza nel task 2 e maggior accuratezza nel task 3.



Mehnert (1998)	31 studenti universitari apprendenti di tedesco, livello intermedio.	qual è l'effetto della durata della pianificazione sulla performance orale?	2 task: (1) task strutturato (istruzioni) (2) task non strutturato (esposizione).  quattro gruppi: (1) nessuna pianificazione (2) 1 minuto (3) 5 minuti (4) 10 minuti.	<i>complessità:</i> - numero di clausole subordinate per T-unit - numero di s-nodes per T-unit <i>accuratezza:</i> - percentuale delle clausole senza errori - errori ogni 100 parole - uso corretto di tre regole sull'ordine delle parole. <i>fluenza:</i> - numero di pause - totale delle pause - length of run - sillabe al minuto <i>densità:</i> - numero degli item lessicali sul totale degli item.	<i>complessità:</i> nessun effetto. <i>accuratezza:</i> la pianificazione e il task strutturato danno maggiore accuratezza. <i>fluenza:</i> La pianificano dà sempre maggiore fluenza. <i>densità:</i> la pianificazione di 10 min nel task strutturato dà maggior densità lessicale.
Ortega (1999)	64 apprendenti avanzati di spagnolo LS.	la pianificazione pre-task migliora complessità, accuratezza e fluenza delle produzioni? Cosa fanno gli apprendenti mentre pianificano?	storia ad immagini  2 condizioni ± pianificazione  10 minuti di pianificazione libera + presa d'appunti.	<i>complessità:</i> - parole per enunciato - type-token ratio <i>accuratezza:</i> - uso corretto dei modificatori nominali - uso corretto degli articoli <i>fluenza:</i> - pruned speech rate  Interviste retrospettive semi-strutturate.	la pianificazione dà miglioramento in tutte le misure, escluso il type-token ratio.

Wigglesworth (2001)	400 apprendenti di inglese LS a diversi livelli di proficiency.	quali effetti hanno le caratteristiche del task e la pianificazione sulla qualità delle produzioni orali?	5 task  pianificazione 5 minuti	<i>valutazioni analitiche:</i> - grammatica - fluenza - coesione - efficacia comunicativa	effetto negativo della pianificazione.
Yuan & Ellis (2003)	42 studenti universitari cinesi laureati in inglese.	quali sono gli effetti della pianificazione pre-task su complessità, accuratezza e fluenza in un task narrativo orale	storia ad immagini - nessuna pianificazione - pianificazione pre-task	<i>complessità:</i> - clausole per T-unit - numero di diverse forme verbali - type-token ratio <i>accuratezza:</i> - clausole senza errori - percentuale di forme verbali corrette <i>fluenza:</i> - sillabe per minuto - pruned sillabe per minuto	<i>complessità:</i> la pianificazione dà più clausole per T-unit. Nessun effetto per le altre misure. <i>accuratezza:</i> nessun effetto. <i>fluenza:</i> la pianificazione dà più sillabe per min.
Skehan & Foster (2005)	61 studenti a livello intermedio	qual è l'effetto l'effetto della pianificazione in due momenti diversi – primi 5 minuti e secondi 5 minuti della performance?	decision-making task (assegnare frasi a scene di crimine)  3 condizioni - nessuna pianificazione - 10 minuti di pianificazione guidata - 10 minuti di pianificazione libera.	<i>complessità:</i> - clausole subordinate per AS-unit <i>accuratezza:</i> - clausole senza errori <i>fluenza:</i> - pause al confine di clausola - pause a metà clausola - pause piene - length or run - riformulazioni - false partenze	<i>complessità:</i> la pianificazione dà più subordinate, ma solo al t1. <i>accuratezza:</i> la pianificazione dà più accuratezza, ma solo al t1. <i>fluenza:</i> la pianificazione dà meno pause.

Tavokoli & Skehan (2005)	80 apprendenti di inglese LS in Iran, livello elementare ed intermedio.	quali sono gli effetti della struttura del task, della pianificazione e della proficiency su complessità, accuratezza e performance in task narrativi orali?	task narrativi con diversi gradi di complessità  5 minuti per guardare le immagini e pianificare il racconto	<i>complessità:</i> - clausole per AS unit <i>accuratezza:</i> - clausole senza errori <i>fluenza:</i> - velocità d'eloquio - silenzio totale - misure di repair (per es. false partenze) - length of run	la pianificazione e il livello danno più complessità, accuratezza e fluenza.
*Elder & Iwashita (2005)	197 adulti apprendenti di inglese L2 in Australia	quali sono gli effetti della pianificazione su narrazioni orali in situazione di testing misurate con valutazioni soggettive e misure analitiche?	2 task narrativi del Test of Spoken English.	<i>valutazione soggettive</i> complessità, accuratezza e fluenza <i>misure analitiche</i> <i>complessità:</i> - clausole per c-unit <i>accuratezza:</i> - percentuale delle clausole senza errori <i>fluenza:</i> - numero di ripetizioni, false partenze, riformulazioni, esitazioni e pause sul totale del tempo.	nessun effetto.
Van Daele <i>et al.</i> (2007)	25 studenti olandesi, 25 studenti inglesi e 25 studenti francesi di 14-16 in Belgio	quali sono gli effetti della pre-pianificazione strategica? Qual è l'effetto della variabile introversione/estroversione? Come si sviluppano complessità, accuratezza e fluenza nelle due LS?	storia ad immagini  4 racconti nelle tre lingue in 6 mesi  gruppo di controllo nativo	<i>complessità:</i> - dipendenti/T-unit - indice di Guiraud <i>accuratezza:</i> - errori lessicali per clausola - errori morfosintattici per clausola <i>fluenza:</i> - velocità d'eloquio (pruned e unpruned speech).	la pianificazione ha effetti su tutte e tre le dimensioni per l'inglese, solo per fluenza e complessità per il francese. Le dimensioni si sviluppano diversamente in inglese e francese. Introversione e ansietà hanno effetti negativi sulla ricchezza lessicale per le due lingue, per l'accuratezza in inglese.



**Tabella 2. Task-based studies, produzione orale e CAF, ripetizione del task**

studio	soggetti	domande di ricerca	metodologia	variabili dipendenti	risultati
Bygate (1996)	1 apprendente	quali sono gli effetti della ripetizione del task sull'uso linguistico?	racconto: cartone animato Tom e Jerry 2 narrazioni	<i>complessità</i> : - subordinazione. - numero di forme verbali. - tipi di connettivi. <i>accuratezza</i> : - percentuale di errori. - collocazioni lessicali. <i>fluenza</i> : - numero di ripetizioni.	La pianificazione dà più accuratezza, maggiore complessità e varietà grammaticale, riduzione del numero di espressioni inappropriate, aumento delle auto-correzioni.
Gass <i>et al.</i> (1999)	104 studenti universitari al quarto semestre di spagnolo.	la ripetizione del task genera un uso più sofisticato del linguaggio? Tale uso più sofisticato si trasmette a un nuovo contesto?	racconto durante la visione video di Mr Bean. 3 gruppi - gruppo stesso contenuto: 3 stesso racconto (intervallo di 2-3 giorni), seguito da un nuovo racconto. - gruppo contenuto diverso: 4 racconti diversi in 4 giornate diverse. - gruppo di controllo - stesso racconto 2 volte.	- valutazioni olistiche della proficiency globale - cambiamenti nell'uso di <i>ser</i> e <i>estar</i> . - lexical sophistication – ratio delle parole complesse sul totale delle parole.	il gruppo stesso contenuto ottiene - risultati migliori tra t1 e t3, ma non in t4 (nuovo task); - migliora nell'uso di <i>ser</i> e <i>estar</i> più del gruppo contenuto diverso; - ottiene una maggior lexical sophistication.
Bygate (1999)	dati da (1998)	diversi task elicitano diverse strutture linguistiche? I diversi task influenzano il processing?	3 task: - 20 domande - discussione - storia	- parole per T-unit - argomenti del verbo - gruppi verbali - 'I think' - frequenza dei nomi	nel task narrativo la ripetizione dà più parole per t-unit è più argomenti del verbo. Nessun effetto per subordinazione

Bygate (2001)	48 apprendenti di inglese L2 residenti negli UK	che effetti produce la ripetizione di uno stesso task su un nuovo task?	2 task - racconto - intervista	<i>complessità:</i> - numero di parole per t-unit <i>accuratezza:</i> - errori per t-unit <i>fluenza:</i> - numero di pause vuote per t-unit	nel task intervista la ripetizione del task genera un aumento della complessità ma porta a una riduzione della fluenza. I risultati non sono statisticamente significativi.
Lynch & McLean (2001)	14 apprendenti di inglese ESP	qual è la relazione tra ripetizione del task, livello di competenza e performance?	presentazione di un poster a 6 visitatori	analisi qualitativa	effetti positivi della ripetizione su accuratezza, pronuncia, lessico e sintassi.
Nemeth & Kormos (2001)	24 ungheresi apprendenti di inglese	qual è l'effetto della ripetizione del task sulla pragmatica dell'argomentazione?	argomentazioni - ripetizione - 1 anno di training	<i>misure pragmatiche:</i> numero di - argomenti - controargomenti - mosse di supporto - mosse di controsupporto	effetto positivo della ripetizione su tutte le misure. Nessun effetto per il training.

**Tabella 3. Task-based studies, produzione orale e CAF, caratteristiche del task**

<b>studio</b>	<b>soggetti</b>	<b>domande di ricerca</b>	<b>metodologia</b>	<b>variabili dipendenti</b>	<b>risultati</b>
Robinson (1995)	8 apprendenti di inglese a livello elementare e intermedio	quali sono gli effetti della dimensione $\pm$ Here-and-Now su complessità, accuratezza e fluenza delle produzioni in un task monologico?	storia ad immagini  2 condizioni - Here-and-Now, con prompt visivo - There-and-Then, senza prompt visivo	<i>complessità:</i> - s-nodes per T-unit - percentuale delle parole lessicali e delle parole grammaticali per utterance <i>accuratezza:</i> - uso target di alcune strutture <i>fluenza:</i> - pause per enunciato - parole per enunciati	nessun effetto significativo sulla complessità. Maggiore accuratezza e complessità lessicale nelle produzioni del task più complesso. Maggiore fluenza per il task più semplice.
Skehan & Foster (1999)	47 adulti a livello intermedio	quali sono gli effetti della struttura del task sulla performance? Qual è l'effetto delle condizioni di processing?	2 storie $\pm$ strutturate video di Mr Bean  2 variabili - struttura del task - carico di processing  4 condizioni - guarda e racconta contemporaneamente - prima storyline, poi guarda e racconta contemporaneamente - prima guarda, poi guarda e racconta contemporaneamente - prima guarda poi racconta	<i>complessità:</i> - subordinate per c-unit <i>accuratezza:</i> - clausole senza errori per clausole <i>fluenza:</i> - fenomeni di esitazione	il carico cognitivo ha effetti sulla complessità. L'accuratezza dipende dall'interazione della struttura del task e carico cognitivo. La fluenza dipende dalla struttura del task.

Robinson (2000)		quali sono gli effetti delle dimensioni $\pm$ reasoning demand sulle produzioni interattive orali?	task narrativo interattivo: riordina le sequenze di una storia ad immagini, racconta la storia al compagno che deve riordinare a sua volta le sue immagini.  2 condizioni $\pm$ dual task $\pm$ reasoning demands	<i>incorporazione:</i> - uptake parziale / turno - uptake esatti / turno	il task più complesso dà maggiore incorporazione delle strutture target.
Niwa (2000)	22 studenti universitari in Giappone	quali sono gli effetti delle dimensioni $\pm$ dual task $\pm$ reasoning demands sulle produzioni? Quali interazione ci sono tra caratteristiche individuali e complessità, accuratezza e fluenza?	task narrativo monologico: riordina le sequenze di una storia ad immagini e racconta la storia al compagno 2 condizioni $\pm$ dual task $\pm$ reasoning demands	<i>complessità:</i> - s-nodes per T-unit - type token ratio <i>accuratezza:</i> - percentuale delle T-unit senza errori <i>fluenza:</i> - parole al secondo - secondi tra le pause (length of run) - parole per pause - parole per T-unit	nessun effetto per complessità e accuratezza. Con l'aumentare della complessità del task, l'intelligenza influisce sulla fluenza. Un maggior reading span of working memory influisce negativamente sulla fluenza.
Robinson (2001b)	studenti universitari	quali sono gli effetti delle dimensioni $\pm$ elementi $\pm$ conoscenze pregresse su complessità, accuratezza e fluenza delle produzioni interattive orali?	map task  2 condizioni: <i>semplice</i> - mappa con pochi elementi - area conosciuta <i>complessa</i> - mappa con molti elementi - area conosciuta	<i>complessità, accuratezza e fluenza</i> come (1995)  <i>interazione:</i> - numero di comprehension check - clarification request	nessun effetto per complessità sintattica e accuratezza. Il task più complesso dà maggior complessità lessicale e fluenza.



*Iwashita <i>et al.</i> (2001)	193 studenti frequentanti i corsi preparatori di inglese per l'accesso all'università	quali sono gli effetti di caratteristiche del task e condizioni di performance su complessità, accuratezza e fluenza?	task narrativo su stimolo visivo.  4 condizioni: ± prospettiva ± immediatezza ± adeguatezza ± pianificazione	<i>misure analitiche complessità:</i> - numero di clausole/ c unit <i>accuratezza:</i> - percentuale di clausole senza errori <i>fluenza:</i> - numero di ripetizioni, false partenze, esitazioni e pause sul totale del tempo parlato  <i>valutazioni soggettive</i> complessità, accuratezza e fluenza	misure analitiche e valutazioni olistiche offrono risultati significativi solo per la dimensione ± immediatezza. le misure analitiche mostrano un aumento dell'accuratezza con l'aumentare della complessità del task, con risultati maggiori per la condizione: narrazione senza prompt visivo.
Gilabert (2007)	48 studenti universitari	quali sono gli effetti di un aumento della complessità del task nelle due dimensioni: ± Here-and-Now, ±Pianificazione time? Come vengono influenzate complessità, accuratezza e fluenza?	2 task narrativi  4 condizioni: (1) planned e H-N (2) unplanned e H-N (3) planned e T-T (4) Unplanned e T-T	<i>complessità:</i> - s-nodes/T-unit - indice di Guiraud <i>accuratezza:</i> - percentuale delle autocorrezioni <i>fluenza:</i> pruned-speech rate	condizione (1): favorisce la fluenza e la ricchezza lessicale; condizione (2): tutte le dimensioni sono influenzate negativamente; condizione (3): favorisce complessità e accuratezza, ma non la fluenza; condizione (4): effetti positivi sul monitoring.
Banerjee <i>et al.</i> (2007)	275 candidati cinesi e spagnoli esame IELTS	quali sono le caratteristiche distintive delle produzioni scritte nelle bande di livello IELTS 3-8? Quali sono gli effetti di task e lingua madre sulla performance?	Task 1 e 2 IELTS Academic writing module	<i>misure analitiche strutture di coesione:</i> - classificazione per tipo - frequenza d'uso <i>ricchezza lessicale:</i> - type/token ratio - densità lessicale - lexical sophistication <i>complessità sintattica:</i> - clausole / T-unit - subordinate / T-unit - tipi di clausole <i>accuratezza grammaticale:</i> - uso di alcune strutture	le misure hanno tutte validità predittiva rispetto ai livelli, tranne complessità. Ai livelli più alti le differenze più significative riguardano il lessico il task 1 produce una maggior densità lessicale. I candidati cinesi ottengono una maggior densità lessicale.

Iwashita <i>et al.</i> (2007)	200 produzioni TOEFL iBT livelli 1-5	qual è la relazione tra feature linguistiche della produzione e livelli?	5 task - indipendenti - integrati	<i>misure analitiche</i> <i>complessità:</i> - clausole / T-unit - subordinate / T-unit - MLU <i>accuratezza:</i> - accuratezza globale - uso di alcune strutture <i>vocabolario:</i> - type/token ratio <i>pronuncia:</i> - target-like syllable <i>fluenza:</i> - numero di pause piene e vuote - tempo di pausa totale - repair - velocità d'eloquio  <i>giudizi olistici</i>	le misure hanno validità predittiva rispetto ai livelli. Fluenza e lessico sono le più significative.
Michel <i>et al.</i> (2007)	44 apprendenti di olandese L2	quali sono gli effetti di un aumento della complessità del task rispetto alle condizioni ± elementi sulla performance orale? Quali sono gli effetti della variazione del task rispetto alla dimensione ± monologico? Ci sono interazioni tra le due condizioni?	decision making task:  2 condizioni ± elementi ± monologico	<i>complessità:</i> - clausole / AS-unit - indice di subordinazione - indice di Guiraud - % parole lessicali <i>accuratezza:</i> - errori per AS-unit - riparazioni per numero di errori - % di self-repair ogni 100 parole <i>fluenza:</i> - sillabe al minuto - numero di pause piene ogni 100 parole	il task dialogico risulta più accurato lessicalmente più complesso e più fluente. Il task monologico ottiene la maggior quantità di repair e una maggior complessità sintattica e lessicale.